

This is a repository copy of *A limited role for unforced internal variability in 20th century warming*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/147408/>

Version: Accepted Version

Article:

Cowtan, Kevin Douglas orcid.org/0000-0002-0189-1437 (2019) A limited role for unforced internal variability in 20th century warming. Journal of Climate. ISSN 0894-8755

<https://doi.org/10.1175/JCLI-D-18-0555.1>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

A limited role for unforced internal variability in 20th century warming.

Karsten Haustein* and Friederike E.L. Otto

Environmental Change Institute, University of Oxford, Oxford, United Kingdom

Victor Venema

Department of Meteorology, University of Bonn, Germany

Peter Jacobs

George Mason University, Department of Environmental Science and Policy, Fairfax, USA

Kevin Cowtan

Department of Chemistry, University of York, York, United Kingdom

Zeke Hausfather

Berkeley Earth, Berkeley, CA, USA; University of California Berkeley, Berkeley, CA, USA

Robert G. Way

Department of Geography and Planning, Queen's University, Kingston, ON, Canada

Bethan White

School of Earth and Atmosphere and Environment, Monash University, Melbourne, Australia

Aneesh Subramanian

17

Scripps Institution of Oceanography, San Diego, CA, USA

18

Andrew P. Schurer

19

School of Geosciences, University of Edinburgh, United Kingdom

20 **Corresponding author address:* Karsten Haustein, Environmental Change Institute, University of

21 Oxford, Oxford, United Kingdom

22 E-mail: karsten.haustein@ouce.ox.ac.uk

ABSTRACT

23 The early 20th century warming (EW; 1910-1945) and the mid-20th century
24 cooling (MC; 1950-1980) have been linked to both, internal variability of the
25 climate system and changes in external radiative forcing. The degree to which
26 either of the two factors contributed to EW and MC, or both, is still debated.
27 Using a two-box impulse response model, we demonstrate that multidecadal
28 ocean variability was unlikely to be the driver of observed changes in global
29 mean surface temperature (GMST) after 1850 A.D. Instead, virtually all (97-
30 98%) of the global low-frequency variability (> 30 years) can be explained
31 by external forcing. We find similarly high percentages of explained vari-
32 ance for inter-hemispheric and land-ocean temperature evolution. Three key
33 aspects are identified which underpin the conclusion of this new study: inho-
34 mogeneous anthropogenic aerosol forcing (AER), biases in the instrumental
35 sea surface temperature (SST) datasets, and inadequate representation of the
36 response to varying forcing factors. Once the spatially heterogeneous nature
37 of AER is accounted for, the MC period is reconcilable with external drivers.
38 SST biases and imprecise forcing responses explain the putative disagreement
39 between models and observations during the EW period. As a consequence,
40 Atlantic Multidecadal Variability (AMV) is found to be primarily controlled
41 by external forcing too. Future attribution studies should account for these im-
42 portant factors when discriminating between externally-forced and internally-
43 generated influences on climate. We argue that AMV must not be used as a
44 regressor and suggest a revised AMV index instead (North Atlantic Variability
45 Index; NAVI). Our associated best estimate for the transient climate response
46 (TCR) is 1.57 K (± 0.70 at the 5-95% confidence level).

47 **1. Introduction**

48 The global temperature evolution over the instrumental period is conventionally attributed to the
49 combination of external forcing and internal variability (Stott et al. 2000; Bindoff et al. 2013; Flato
50 et al. 2013). Virtually all of the warming since 1950 is attributed to human influences (Stocker
51 et al. 2013; Jones et al. 2013, 2016; Ribes et al. 2017). Yet due to the loosely constrained nature
52 of magnitude and evolution of AER, there continues to be a fierce debate about the cause of
53 multidecadal GMST fluctuations present in the instrumental record (Shiogama et al. 2006; Booth
54 et al. 2012; Zhang et al. 2013; Thompson et al. 2015). Most prominently, the origin of the EW
55 and MC periods, thought to be linked with North Atlantic (NA) ocean variability and commonly
56 expressed in terms of AMV (Delworth and Mann 2000; Knight et al. 2005, 2006), is still hotly
57 contested because of the difficulties to disentangle the contributions from internal and external
58 drivers at different timescales (Brönnimann 2009; Mann et al. 2014; Zhang et al. 2016; Clement
59 et al. 2016; Vecchi et al. 2017; Sutton et al. 2017; Hegerl et al. 2018).

60 Conventionally, the AMV has predominantly been attributed to internal ocean variability, which
61 in turn has been linked to changes in the Atlantic Meridional Overturning Circulation (AMOC)
62 as a deep ocean driving mechanism on multidecadal timescales (Zhang and Wang 2013; Yeager
63 and Robson 2017). While stochastic atmospheric flux forcing is thought to influence SSTs on
64 shorter timescales (Roberts et al. 2013; Duchez et al. 2016; Josey et al. 2019), associated with
65 changes in the North Atlantic Oscillation (NAO) index (Hurrell and Deser 2009), the prevailing
66 view regarding NA SST changes on longer timescales is that large internal variations are superim-
67 posed on the anthropogenic warming trend. However, in recent years, external forcing has been
68 shown to contribute to multidecadal swings in the AMV region (Otterå et al. 2010; Murphy et al.
69 2017; Bellucci et al. 2017), suggesting a reduced role for internal ocean dynamics. Changes in

70 AER (Booth et al. 2012; Bellomo et al. 2018) as well as periods of strong volcanic activity (Iwi
71 et al. 2012; Knudsen et al. 2014; Pausata et al. 2015; Swingedouw et al. 2017) have been linked
72 to these changes. Also, it has been demonstrated that AMV-like SST pattern can be reproduced in
73 slab-ocean experiments (Clement et al. 2015, 2016; Bellomo et al. 2018). Hence internally gen-
74 erated low-frequency GMST variations are increasingly thought to play only a smaller role, with
75 Pacific ocean variability to be more recognised as a pacemaker for global temperature (Schurer
76 et al. 2015; Dong and McPhaden 2017).

77 While there is no debate about the existence of aerosol-related dimming and brightening (Wild
78 et al. 2007; Wild 2009) due to a huge array of supporting data from observations (Boers et al.
79 2017; Dumitrescu et al. 2017; Manara et al. 2017) and modelling (Shindell et al. 2013; Wilcox
80 et al. 2013; Rotstayn et al. 2015; Dallafior et al. 2016; Chung and Soden 2017), its impact on the
81 AMV is less certain. Many studies do not (Huss et al. 2010; Chylek et al. 2014) or insufficiently
82 (Ting et al. 2009; Zhang et al. 2013) incorporate or acknowledge AER, which potentially leads to
83 misattribution of cause (Zhang et al. 2016; O'Reilly et al. 2016) and effect (Chylek et al. 2009;
84 Wyatt et al. 2012; Tung and Zhou 2013; Pasini et al. 2016; Levine et al. 2018). Arguments for
85 the presence of an internally-generated AMV based on ostensible pseudo-oscillatory behaviour in
86 instrumental, proxy, or model data are unconvincing (Singh et al. 2018), and it is noted that such
87 behaviour can arise from statistical artefacts alone (Vincze and János 2011; Cane et al. 2017).
88 Regression-based methods are thereby particularly susceptible to conflating internal variability
89 with forced responses because of strong covariance between the predictors (Mann et al. 2014;
90 Stolpe et al. 2017), yet studies that use the AMV as regressor or explanatory factor continue to
91 be published despite the lack of an unequivocal physical underpinning (Lewis and Curry 2018;
92 Rypdal 2018; Shen et al. 2018; Zhang et al. 2018; Folland et al. 2018).

93 We argue that any attribution exercise that does not sufficiently account for the spatio-temporal
94 AER changes will invariably produce unreliable and erroneous results. Incorporating now better
95 quantifiable biases in the instrumental SST record, we demonstrate that a carefully designed anal-
96 ysis (that avoids overfitting) yields a surprisingly high level of agreement between our model and
97 observations without the need to infer additional unexplained internal variability. We endeavour
98 to highlight the pitfalls associated with attributing and identify the shortcomings in representing
99 the externally forced temperature responses.

100 Since attempts to estimate the magnitude of internal variability by means simple climate models
101 are plagued from dissatisfying low correlations with observations (Aldrin et al. 2012; Skeie et al.
102 2014), here we use a refined two-box impulse response model framework which accounts for fast
103 and slow responses to forcing perturbations in the climate system. To constrain the complexity of
104 the model, we introduce a novel TCR adjustment factor for different forcing agents that is gov-
105 erned by robust physical factors. Apart from Northern Hemisphere (NHem) and GMST (Global),
106 we also analyse Southern Hemisphere (SHem), land surface air temperature (Land) and SSTs
107 (Ocean), expanding on previous GMST-only analyses (Mann et al. 2014; Dong and McPhaden
108 2017) to better understand the impact of radiative forcing changes on surface temperatures. We
109 recommend all impulse response or energy balance model studies use land, ocean, and hemi-
110 spheric temperature records with our dedicated set of model parameters as separate benchmark
111 tests to robustly evaluate model performance.

112 **2. Radiative forcing and observational data**

113 We use the latest well-mixed greenhouse gas (WMGHG) radiative forcing (Etminan et al. 2016;
114 Meinshausen et al. 2017) and the gridded aerosol community emission dataset (CEDS) (Hoesly
115 et al. 2017), including sulphur dioxide (SO_2), ammonia (NH_3), black carbon (BC), and organic

116 carbon (OC). For solar forcing, we use sunspot numbers from the Greenwich Royal Observatory
117 (Wilson and Hathaway 2006), scaled to solar forcing according to Dewitte and Nevens (2016).
118 Stratospheric aerosol optical depth (AOD) data from explosive volcanic eruptions (Crowley et al.
119 2008; Crowley and Unterman 2013) are scaled to match NASA-GISS volcanic forcing data (Sato
120 et al. 1993), and updated to include recent smaller eruptions (Vernier et al. 2011; Solomon et al.
121 2011; Arfeuille et al. 2014; Schmidt et al. 2018). Fig. 1a shows our revised forcing estimates.

122 The global direct radiative forcing for each aerosol component (SO_2 , NH_3 , BC, OC) is derived
123 by scaling the current emissions to the AR5-forcing estimate for 2011 (Myhre et al. 2013; Stocker
124 et al. 2013). Using BC emissions over North America, we account for enhanced Arctic warming
125 during the first half of the 20th century (Johannessen et al. 2004; McConnell et al. 2007; Mc-
126 Connell and Edwards 2008; Suo et al. 2013) (orange shading in Fig. 1b). The indirect forcing
127 of -0.45W/m^2 is mostly a function of SO_2 (90%; 10% for OC). While considerable uncertainty
128 regarding aerosol-cloud effects exist (Carslaw et al. 2013; Regayre et al. 2014; Nazarenko et al.
129 2017; Lohmann 2017), the best estimate for indirect AER in AR5 has not been fundamentally
130 challenged since. Together with the direct effects, we obtain a total AER of $\sim -0.55\text{W/m}^2$ in ac-
131 cordance with AR5 (Fig. 1b), which is set to -0.75W/m^2 *pseudo-effective* global aerosol radiative
132 forcing (ERF) in our response model framework (Fig. 1b, c). The total ERF estimate is guided
133 by a recent review by Forest (2018), which is slightly lower than the best estimate for ERF of
134 -0.9W/m^2 published in AR5. We note that other recent research has also suggested that AER ERF
135 might be lower (Stevens 2015; Myhre et al. 2017), essentially reflecting arguments for stronger BC
136 warming effects (Bond et al. 2013; Myhre and Samset 2015) and less cooling due to noticeable
137 SO_2 reductions in China since 2006 (Smith et al. 2011; Klimont et al. 2013).

138 We use Berkeley Earth Land/Ocean (BE) (Rohde et al. 2012), HadCRUT4-Cowtan/Way
139 (Cru4CW) (Cowtan and Way 2014; Cowtan et al. 2015), HadISST2 (Titchner and Rayner 2014;

Kennedy et al. 2017) and OSTIA (Donlon et al. 2012) as observational data. We note that there are indications that the land datasets may still underestimate warming in northern areas (Wang et al. 2017; Way et al. 2017). Since Cru4CW uses HadSST3 data (Kennedy et al. 2011) over oceans, we developed an additional composite product with Cru4CW over land and HadISST2 (1850-1985; preliminary release available only until 2010) and OSTIA data (1986-now; calibration period 1986-2005) over ocean to reflect the full range of available SST products (hereinafter referred to as HadOST). To obtain land and ocean proxies, ocean points that are covered with sea ice are treated as land points. The sea ice extent to generate the ice mask is taken from HadISST2 and OSTIA. The same mask is applied to Cru4CW and BE.

Due to continuous problems in currently available SST datasets, mainly manifest as warm bias as a result of changing SST sampling methods (from bucket to engine-room intake measurements) during World War II, associated with changing fleet composition (Karl et al. 2015; Hansen et al. 2016; Kent et al. 2017), Cowtan et al. (2017) have recently proposed a novel method to address the WWII bias using island and coastal weather stations only. Inspired by the idea, we replicate their analysis with a slightly simplified methodology. We use a mask where grid boxes over land (adjacent to ocean) and over ocean (along coastlines) are selected, including islands. The global average of all such subsampled ocean grid boxes establishes our new SST proxy. The two coastal time series are scaled to match the 1980-2016 global SST trend (see Cowtan et al. (2017) for details on the scaling method) as it is deemed the most reliable period in the marine instrumental record (Rahmstorf et al. 2017).

The results are shown in Fig. 1d for HadOST and BE and in Fig. 1e for ERSSTv4 and GHCNv3 as used in GISTEMP (Hansen et al. 2010). The scaling factors are provided in the figure legend. In both cases, the two coastal records (derived from HadISST2 and ERSSTv4) show excellent agreement during the calibration period. As expected, the land scaling factor is lower in agreement

164 with amplified warming trends over land. The land and ocean proxies agree after 1920 and show
165 only minor deviations before 1920 (Fig. 1d). The HadOST proxies (land and ocean), suggest that
166 HadISST2 is reliable with marginal biases between 1880-1940. We find much less agreement
167 between GHCNv3 and ERSSTv4 before 1980. Our analysis further suggests that ERSSTv4 has
168 a substantial cold bias between 1900-80 as well as a spurious warm bias during WWII (Fig. 1e).
169 While by no means perfect, this straight-forward analysis is at least indicative that SSTs in general
170 and ERSST in particular (versions 4 and 5 are almost identical throughout the period of coverage)
171 are still impacted by substantial unresolved inhomogeneities. In our main analysis we discard
172 GISTEMP and apply the following correction factors to HadOST, Cru4CW and BE during four of
173 the WWII years (1942-45): NHem = -0.04°C , Global = -0.08°C , SHem = -0.12°C and Ocean =
174 -0.18°C . The remaining years in the time series remain unchanged. We discuss the implications in
175 section 4.

176 Finally, we use historical climate simulations from the Coupled Model Intercomparison Project
177 (CMIP5) (Taylor et al. 2012) and an ensemble of the UK MetOffice HadCM3 model (Euro500)
178 (Schurer et al. 2014) to estimate warming ratios and multidecadal internal variability.

179 **3. Impulse response model and uncertainty**

180 Following the method introduced in earlier work (Otto et al. 2015; Haustein et al. 2017),
181 (vaguely similar to the analysis presented in Lean and Rind (2008) and Lean (2018)) we employ
182 a two-box impulse response model framework that accounts for fast and slow temperature (T)
183 changes in response to external forcing factors (*comp*: WMGHGs, anthropogenic aerosols (AER)
184 and volcanic eruptions (VOL)). The fast component can be associated with the ocean mixed layer
185 response whereas the slow component approximates the response of the deep ocean (Li and Jarvis
186 2009):

$$\frac{dT_j}{dt} = \frac{q_j \cdot F - T_j}{d_j}; T_{comp} = \sum_{j=1}^2 T_j \quad (1)$$

$$\begin{aligned} TCR_{comp} &= F_{2xCO_2} \cdot \left(q_1 \cdot \left(1 - \frac{d_1}{70} \left(1 - e^{-\frac{70}{d_1}} \right) \right) \right) \\ &+ q_2 \cdot \left(1 - \frac{d_2}{70} \left(1 - e^{-\frac{70}{d_2}} \right) \right) \end{aligned} \quad (2)$$

$$ECS = F_{2xCO_2} \cdot (q_1 + q_2) \quad (3)$$

More details can be found in Millar et al. (2017). The forcing due to doubling of CO₂ (F_{2xCO_2}) is 3.71 W/m². The factor q_j (integrated contribution for response j) can be determined using Equation (2) and (3) with a defined set of values for TCR and ECS. Our chosen TCR range encompasses values from 1.1-2.1K, with an associated ECS range of 2.0-4.0K, in line with IPCC AR5 estimates (Stocker et al. 2013). As TCR/ECS-ratios derived from observational data are plagued by a variety of shortcomings (Armour 2017; Proistosescu and Huybers 2017; Marvel et al. 2018), we apply the CMIP5 mean of ~ 0.53 as our central TCR/ECS-ratio estimate, supported, for example, by a reasonably good match of measured and simulated ocean heat uptake (Cheng et al. 2016). The associated adjustment factors for NHem, SHem, Land and Ocean as well as for AER and VOL forcing are introduced below.

The slow response time (d_2) is taken from Geoffroy et al. (2013a) (320 years), which included deep ocean feedbacks in contrast to accompanying work (Geoffroy et al. 2013b). Given that the fast response time (d_1) of 4 years suggested in the same study (Geoffroy et al. 2013a) relies on estimates from GCM simulations, we follow the approach presented in Rypdal (2012) and double d_1 to 8 years, which is in line with coefficients presented in Boucher and Reddy (2008) based on idealised simulations undertaken with the HadCM3 model. It is argued that observed

temperatures show a prolonged/delayed response due to mediating effects intrinsic to our climate system (Emile-Geay et al. 2008; Santer et al. 2014; McGregor et al. 2015) which may be less well represented in many GCMs (Le 2017). These estimates of d_1 and d_2 yield the highest correlation with observations.

As far as the response to AER is concerned, Shindell (2014) and Marvel et al. (2016) have highlighted the importance of different hemispheric treatment of the heterogeneous aerosol load. The conceptual idea is to have an enhanced TCR for AER due to its preponderance over land as a result of the skewed spatial distribution of aerosols. Differential heat capacities over land and ocean (and therefore implicitly the hemispheres) lead to considerably different response times over land and ocean, associated with inhomogeneous hemispheric warming rates that are mediated by cross-equatorial energy transports (Loeb et al. 2016; Stephens et al. 2016) for all forcing agents. Having said that, aerosols are transported over vast distances (Uno et al. 2009; Schulz et al. 2012), affecting oceans directly (due to albedo effect) and indirectly as well (due to cloud effects, particularly over formerly pristine areas), despite very low direct emissions over oceans mainly from ship exhaust (Kunkel et al. 2013; Shindell et al. 2013). Therefore, Ocean aerosol emissions are not a suitable proxy for the associated ocean temperature response. To remedy the problem, the inter-hemispheric exchange of aerosol-induced temperature responses has to be accounted for appropriately using coupling factors (introduced below).

The differential warming requires dedicated TCR calibration factors for the WMGHG, VOL and AER induced temperature responses. To obtain a plausible and robust set of such calibration factors, we use observed Transient Warming Ratios (TWR) between NHem and SHem as well as Land and Ocean. In Fig. 2, the temperature responses to total anthropogenic (a, b), WMGHG (c, d), AER (e, f) and VOL (g, h) are shown. Decadally averaged warming ratios are provided above

226 or under each graph. All data are low-pass filtered with an smoothing radius of 5 years. The TWR
227 is obtained during the 30 year period of strongest transient warming.

228 Given that TWRs for WMGHG, AER and VOL can only be inferred from GCMs, we apply a
229 scaling factor which represents the difference between observed and all-forcing TWRs. Assuming
230 that the observed TWR (red shaded area in Fig. 2a, b) is our target ratio, the responses in HadCM3
231 are scaled accordingly. HadCM3 is used because it provides a small ensemble of simulations
232 (mainly drawn from the Euro500 experiment) which is consistent across the experiments. The
233 TWR in the historical HadCM3 ensemble is 1.7, compared to 2.8 in HadOST (Fig. 2a). Hence a
234 scaling factor of ~ 1.6 is applied to the TWR deduced from the WMGHG and AER ensemble of the
235 same model in order to correct for the underestimated TWR in the historical HadCM3 simulations.
236 The resulting inferred TWR (hereinafter referred to as TWRD; D = diagnosed), which is then used
237 in the response model, is provided in the boxes at the bottom of Fig. 2.

238 Since the bulk of the VOL response takes place on the fast timescale (1-10 years) and thus differ
239 from WMGHG related responses (Ding et al. 2014), we refrain from scaling and use the TWR
240 from HadCM3 directly (consistent with above-mentioned findings in Boucher and Reddy (2008)
241 regarding HadCM3's fast response time). Note that the VOL responses in Fig. 2g, h are shown
242 for the full 1500-1999 period in contrast to the shorter 1850-1999 (1850-2017) period for all other
243 scenarios.

244 In addition, since we do not know the resulting warming ratios in the response model a priori
245 when we impose the inferred TWRD, we compare them with the posteriori TWRs (hereinafter
246 referred to as TWRE; E = estimated) in order to validate our approach. We find that, for example,
247 the TWRD for WMGHGs (TCR of 2.65K over Land and 1.11K over Ocean) of ~ 2.4 results in a
248 TWRE of ~ 2.2 (see Fig. 2d). We therefore argue that our method is reasonably well constrained
249 to provide a robust answer.

250 All TCR calibration factors based on the deduced TWRDs (and shown at the bottom of Fig. 2)
 251 are summarised in the upper box in Fig. 3. We would like to point out that these calibration factors
 252 modulate the TCR/ECS ratio and are used for the full range of TCR and ECS values, respectively,
 253 not only the best estimate. The latter is provided at the top of Fig. 3 as well, together with the TCR
 254 for AER effective forcing which is $\sim 40\%$ higher (best estimate = $2.2K$) than that of WMGHGs
 255 (best estimate = $1.6K$), consistent with findings in Rotstayn et al. (2015) and in pursuit to reflect
 256 the higher aerosols load over land.

257 To estimate the TCR calibration factor for AER, hemispheric and land-ocean coupling factors
 258 need to be determined. They reflect the above-mentioned fact that inter-hemispheric energy ex-
 259 changes in response to the heterogeneous distribution of AER need to be balanced. Conveniently,
 260 the coupling factors are an emergent property and as such a function of the hemispheric area
 261 weighting factors, which are strictly interlinked and hence constrained as follows (example for
 262 WMGHGs):

$$\begin{aligned}
 T_{GHG}^{Global} &= 0.5 \cdot T_{GHG}^{NHem} + 0.5 \cdot T_{GHG}^{SHem} \\
 &= 0.32 \cdot T_{GHG}^{Land} + 0.68 \cdot T_{GHG}^{Ocean}
 \end{aligned} \tag{4}$$

263 Note that the Land fraction is marginally $>30\%$ because areas covered with sea ice are treated
 264 as land throughout the analysis. Apart from the area-weighted constraint, the coupling factors
 265 are also dependent on the emission ratio, i.e. the ratio between the hemispheric (and land/ocean)
 266 and the total global aerosol emission strength, which in turn determines the appropriate fractional
 267 contribution to match the inferred AER-TWRD (see Appendix A for more details). The resulting
 268 coupling factors are 3.9 (ratio of 1.47 and 0.38, which corresponds to 85% NHem and 15% SHem
 269 AER contribution for NHem AER and vice versa for SHem AER) and 2.1 (ratio of 1.46 and 0.7,

270 which corresponds to 70% Land and 30% Ocean AER contribution for Land AER and vice versa
271 for Ocean AER). These factors are also provided in the bottom box of Fig. 3, together with all
272 other parameters used in the response model. Global temperature trends for the 1978-2017 period
273 in HadOST (a), CMIP5 (b) and HadCM3 (c) are also shown in Fig. 3. The spatial distribution
274 of the trend highlights why observed and modelled TWR do not agree, which is primarily caused
275 by delayed southern ocean warming (Armour et al. 2016), and partly by an accelerated Arctic
276 amplification (Serreze and Barry 2011). Both physical processes are not satisfactorily reproduced
277 in most GCMs.

278 Lastly, as apparent from the discrepancy between the AER factor provided at the bottom of Fig. 2
279 (3.5 and 2.4 for NHem/SHem and Land/Ocean, resp) and that shown in the top box of Fig. 3 (5.1
280 and 2.9 for NHem/SHem and Land/Ocean, resp), we increased the inferred AER-TWRD slightly.
281 While the adjustment of the AER-TWRD does not change our conclusions (see Fig. S1 and S2 for
282 the same result without AER-TWRD tuning), it does lead to better agreement between HadOST
283 and the response model during the period of strongest AER cooling between 1960-80. Given
284 that the HadCM3 AER ensemble is not a strict AER-only simulation rather than the difference
285 between the *allforcing* and a non-aerosol ensemble of HadCM3, the results likely do not reflect
286 the full extent of the aerosol-induced TWR. Therefore we think it is a defensible decision and well
287 within the realm of the uncertainty of our AER-TWR estimate. The resulting AER timeseries is
288 shown in Fig. 1c.

289 For the uncertainty analysis, response model, radiative forcing and internal variability uncer-
290 tainty is considered. Apart from the TCR (1.1...2.1K) and ECS (2.0...4.0K) range, we also include
291 a range of fast response times (3...13 years) in our response model uncertainty estimate. For the
292 forcing, 200 total radiative forcing realisations are used (Forster et al. 2013) and converted into
293 response model temperature equivalents to estimate the associated error range. The resulting σ

(32-68th percentiles) of the fractional uncertainties is shown in Fig. 4a (response model error in green and radiative forcing error in blue). If we assume that potential internally generated, low-frequency variability adds linearly to the externally forced response, we need an estimate of (modelled) unforced multidecadal variability. As introduced in Haustein et al. (2017), we use equidistant intervals of selected CMIP5 pre-industrial control simulations that do not drift (Knutson et al. 2013) and possess a similar range of unforced variability as our response model based estimate of the residual observational variability. In Fig. 4b, the low-pass filtered residuals for HadOST, Cru4CW and BE between 1850-2017 and low-pass filtered sample intervals of 168 years from selected CMIP5 models are shown together with their standard deviation (σ). The obtained 5-95th percentiles of their internal variability span $\pm 0.17^\circ\text{C}$ ($\bar{\sigma}=0.1^\circ\text{C}$ and $\bar{\sigma}^2=0.01^\circ\text{C}^2$) as shown in Fig. 4a in grey.

We note that there is additional parameter uncertainty, which is not fully included here as it is difficult to objectively constrain the upper and lower bounds of the respective parameters. In order to rectify this problem, in Fig. S3, we have plotted the response model results for a set of reasonable model parameters, including aerosol sensitivity, TWRs, coupling strength, TCR efficiency, varying SOL and VOL forcing, as well as high and low AER ERF (dashed line for -0.5 and -1.0 W/m²). The resulting uncertainty is small compared to the total uncertainty, which is dominated by the forcing uncertainty. Hence we conclude that the our results are insensitive to the parameter choices, even if our observationally constrained estimates were biased.

4. Model performance and evolution

In Fig. 5, the response model results for Land (a; brown), NHem (b; red) and Global (c; green), Ocean (d; purple) and SHem (e; blue) are shown. The central *allforcing* temperature response estimate is shown as the bold line in the lower graph in each panel, while thin lines indicate slightly

317 higher/lower alternative TCR estimates as indicated at the right hand side (1.2...2.0K). The 5-
318 95th percentiles and the inter-quantile (25-75th) uncertainties are added as shaded grey contours.
319 The low-pass filtered (30 year smoothing radius) instrumental data from HadOST (dark green),
320 Cru4CW (yellow) and BE (black) are shown for comparison, including the WWII correction in-
321 troduced in section 2. Before filtering, the influence of ENSO (Deser et al. 2012) is removed from
322 the observational timeseries in order to minimise short-term noise (Stuecker et al. 2015), following
323 the multiple regression approach of Foster and Rahmstorf (2011).

324 Conversely, in the upper graphs in Fig. 5a - 5e we have added ENSO variability to the response
325 model results by scaling the multivariate ENSO index (MEI) (Wolter and Timlin 1998) for each do-
326 main and applying the lag coefficient obtained from the multiple regression. Other than the WWII
327 bias correction, the observational data in the upper graphs show the annual mean temperatures. On
328 the top left in each panel, the explained variance (R^2) for non-ENSO corrected, model-adjusted
329 (MEI), and observation-adjusted correlations between model and the observational datasets are
330 shown. The correlations are based on the low-pass filtered timeseries (30 year smoothing radius).
331 To avoid problems due to autocorrelation, the associated non-filtered R^2 between Global HadOST
332 and model-adjusted (MEI) timeseries is 0.935 (not shown; 0.92 for Cru4CW and 0.912 for BE).
333 We would like to highlight that our R^2 for HadOST exceeds the explained variance found in Ryp-
334 dal (2018) and Folland et al. (2018), without the need to invoke any contribution of the contentious
335 AMV.

336 We find excellent agreement between our response model and observations in all three time-
337 series. NHem and Land are well reproduced over the entire duration of the instrumental period,
338 including the EW and the MC periods (Fig. 5a, 5b). SHem and Ocean are similarly well re-
339 produced, with notable deviations before and after WWII when compared with Cru4CW or BE
340 (Fig. 5d, 5e). Using HadOST, the SHem and Ocean model results can be almost entirely reconciled

341 with observations (Fig. 5d). HadOST and Ocean only start to diverge before 1900. But overall,
342 the Global results (Fig. 5c) leave little room (of the order of $\sim 0.1^{\circ}\text{C}$) for unforced low-frequency
343 temperature variations.

344 Before we investigate other notable excursions in light of the role of unforced Pacific and At-
345 lantic ocean variability, in Fig. 6 the evolution of the response model for all five domains is shown.
346 The top graph in each panel shows the response model result using WMGHG and aerosol forc-
347 ing based on IPCC AR5 (Meinshausen et al. 2011), extrapolated to 2017, volcanic forcing from
348 NASA-GISS (Sato et al. 1993) updated to 2017 and a fast response time of 4 years. As such, it
349 corresponds to the results published in Haustein et al. (2017). The middle graph in each panel is
350 using our slightly modified VOL and solar forcing and a fast response time of 8 years. All what
351 is otherwise different compared to our final response model result as shown in the lower graph
352 in each panel is AER. The results based on the new CEDS AER show significant improvements
353 in each domain, resolving most of the discrepancies associated with the EW and MC period. As
354 far as EW is concerned, the improved response model performance is partly linked with the SST
355 bias correction during WWII which is only applied in the lower graph in Fig. 6. Accordingly,
356 the warming spike particularly over Ocean (Fig. 6d) and SHem (Fig. 6e) disappears, leading to a
357 visibly better agreement between model and observations.

358 With the current AER lowered by $>10\%$ (Fig. 1a), here we briefly explore the implications for
359 TCR, including a cautionary remark regarding the lack of robustness when estimating ECS. In
360 Fig. 5, the TCR range from 1.2-2.0K is indicated with our best estimate using a TCR of 1.6K
361 (bold lines). Based on linear regression between HadOST and the Global response model re-
362 sult, our most precise TCR estimate is 1.57K with an associated inter-decile uncertainty range of
363 0.87-2.27K (10-90th percentiles). This is in good agreement with other recent work (Richardson
364 et al. 2016), despite the lower AER estimate. While others have suggested that TCR might be

time-dependent (Gregory et al. 2015), our results do not provide evidence for a change over the instrumental period.

With the TCR/ECS-ratio held constant, ECS is tied to TCR by construction in our analysis (3.0K with an associated inter-decile uncertainty range of 1.7-4.3K). Nonetheless, it is instructive to investigate the impact of different ECS values upon the model results when the TCR/ECS ratio is permitted to vary. As shown in the lower graphs of Fig. 6 where we have added the response model result for the low-end (2.0K) and high-end (4.0K) ECS range, neither of the two estimates provides sufficient guidance as to which ECS value is more likely to be correct. The small range of possible outcomes severely hampers a robust ECS estimation. We therefore agree with others (Armour 2017; Proistosescu and Huybers 2017; Marvel et al. 2018) who found that ECS cannot reliably be inferred from historical observations alone, and recommend caution as ECS is easily conflated with the Effective Climate Sensitivity, the latter of which is likely to be lower (Knutti et al. 2017; Andrews and Webb 2018). Hence such attempts (Aldrin et al. 2012; Otto et al. 2013; Skeie et al. 2014; Lewis and Curry 2015; Mauritsen and Pincus 2017; Lewis and Curry 2018) should be viewed with extreme skepticism.

5. Role of unforced Pacific ocean variability

Returning to Fig. 5, here we assess a few noteworthy remaining excursions that are arguably related to unforced internal variability. To facilitate quantifying those excursions, in Fig. 7 the residuals between the HadOST and response model temperature timeseries are plotted for the five domains. In Fig. 7b, the low-pass filtered MEI evolution is provided (black line). Cru4CW and BE are shown in Fig. 7e for completeness.

We note that MEI shows signs of multidecadal variability, which is linked to the Pacific Decadal Variability (PDV) index (Newman et al. 2016; Henley 2017). Whether or not the unique behaviour

388 of the North Pacific variability (Williams et al. 2017; Kohyama and Hartmann 2017) and the
389 associated observed strengthening of the Walker circulation (L'Heureux et al. 2013; McGregor
390 et al. 2014; de Boissésou et al. 2014; Ma and Zhou 2016; Kajtar et al. 2017) are unforced or partly
391 caused by changes in WMGHG (DiNezio et al. 2012; Xiang et al. 2014; Cai et al. 2015), AER
392 (Dong et al. 2014; Takahashi and Watanabe 2016), or VOL (Emile-Geay et al. 2008; Le 2017) is a
393 matter of intense debate and beyond the scope of this paper. However, the residuals as well as the
394 explained variabilities provided in Fig. 5 suggest that low-frequency ENSO variability has little
395 bearing on the outcome of our response model results. Merely the timing of the modern warming
396 is slightly better aligned with observations when MEI rather than NINO3.4 (Trenberth 1997) is
397 used (not shown), which is indicative of a minor role for additional decadal PDV impacts indeed.

398 Modelled Land (Fig. 7d) shows only a few peaks that are not explained by ENSO (e.g. 1884,
399 1913, 1939, 1949, 1980, 1991, 2010). Such excursions should be expected given the large standard
400 deviation over land due to the stochastic nature of continental interannual variability (Mahlstein
401 et al. 2012). There are a few years between 1950-60 which appear to be cooler than the response
402 model suggests, but since no such deviation shows up over Ocean (see Fig. 7f), it might be related
403 to European aerosol emissions (Persad and Caldeira 2018).

404 The positive residual after 2000 (also visible in the NHem residual in Fig. 7a) is perhaps more
405 interesting as it relates to the infamously dubbed "hiatus" period in the wake of the strong El
406 Niño in 1997/98. While primarily caused by a clustering of La Niña events around 2010 (Kosaka
407 and Xie 2013; England et al. 2014; Schurer et al. 2015; Dong and McPhaden 2017), upon closer
408 inspection another feature stands out. There has been a succession of anomalously cold years
409 between 2010-2013, which is exclusively linked with boreal winter. More precise, this period
410 is linked with extremely cold Eurasian winters (Cohen et al. 2012) which may or may not have
411 been assisted by forced atmospheric circulation changes in response to declining sea ice (Tang

et al. 2013; Cohen et al. 2014; Overland 2016; Francis 2017; Hay et al. 2018). But other than that, SHem (Fig. 7c) and Ocean (Fig. 7f) residuals are inconspicuously smooth and only diverge before 1900 as outlined above already. Overall, our results support previous work that has shown that using updated external radiative forcing (Huber and Knutti 2014; Schmidt et al. 2014) and accounting for ENSO-related variability explains the so-called "hiatus". We refer to Medhaug et al. (2017) for a comprehensive review of the unprecedented flurry of publications on the subject. That said, despite being less sensitive to small changes near the endpoints compared to higher degree polynomial fits, we caution that the lowess smoother is still susceptible to overestimating trend changes at the beginning and end of the time series.

With explained variabilities $\sim 98\%$ for HadOST for the Land (Fig. 5a), NHem (Fig. 5b) and Global (Fig. 5c) response model results, we conclude that almost all low-frequency variability is explained by external forcing factors independent of ENSO. The Ocean (Fig. 5d) and SHem (Fig. 5e) results reveal similar explanatory skill with explained variabilities between 93-95%. Interestingly, BE shows lower correlation factors than Cru4CW over Ocean (even more so over SHem), despite their common use of HadSST3. Thus, differences in data processing alone can explain much of the discrepancies. The fact that HadOST not only fares considerably better in terms of correlation, but also performs best regarding the coastal proxy analysis (Fig. 1d), justifies its inclusion in our analysis. However, more work needs to be done to reconcile the differences between the available SST products and to reduce associated biases (Davis et al. 2019). In Appendix B, we briefly analyse the spatio-temporal characteristics of those products with regard to decadal means.

6. Role of unforced Atlantic ocean variability

While the accurate reproduction of the EW and MC period in our response model framework does not require multidecadal temperature variability to attribute to the ostensible AMV, we do not dispute the existence of internal variability associated with AMOC variations. Therefore here we aim at quantifying the AMOC's role in setting NHem temperatures and its relation to the AMV. In order to facilitate the assessment, we would like to propose a more adequate, straight-forward and intuitive definition of the AMV index itself. Rather than using the standard definition (Delworth and Mann 2000) or an improved definition thereof (van Oldenborgh et al. 2009), we define the AMV as average SST at 25-60°N and 7-75°W (red box in Fig. 8d) minus NHem temperature. The resulting revised timeseries is shown in Fig. 7a (bold black line).

The revised AMV index (which we more appropriately propose to be named *North Atlantic Variability Index* or NAVI) is essentially reflecting and reliably mirroring the long-term AMOC decline in response to anthropogenic warming. The unprecedented dip around 2015 is associated with the continued advection of very cold air of Arctic origin over the Canadian archipelago region during the winters of 2014/15 and 2015/16. Atmospheric forcing has been recognised to drive short-term AMOC variability (Roberts et al. 2013; Ducheze et al. 2016) as opposed to gradual changes in sea ice cover (Sévellec et al. 2017), temperature, salinity and pressure gradients that eventually cause the slower long-term AMOC changes that are indeed already detectable Rahmstorf et al. (2015) and concomitant with the well-known Atlantic Warming Hole (AWH) (Menary and Wood 2018). Arguably, asymmetric land-ocean warming is a more mundane explanation for the colder NA region relative to NHem, as it is physically consistent with a transient warming scenario, but the slow pace of the NAVI decline suggests a contributing role for AMOC.

455 In order to qualitatively explore the role of longer-term effects associated with low-frequency
456 modes of variability, we have conducted a simple correlation analysis. In Fig. 8, we have plotted
457 the spatial map of correlation coefficients between Global and NHem timeseries obtained from the
458 response model versus global observations (HadOST). The correlation between the older, slightly
459 more advanced AMV index (van Oldenborgh et al. 2009) and HadOST is provided as well (Fig. 8d,
460 g, k). We notice that the AWH in the subpolar NA region appears uncorrelated with the forcing
461 timeseries (Fig. 8a, b, e), regardless whether we use the Global or NHem timeseries. Another
462 noteworthy feature is the accompanying anti-correlation between the AMV index and most world
463 regions.

464 Since we are not aware of a robust mechanism that would cause multidecadal AWH variability
465 as opposed to a steady decline, in the following we test three potential reasons for why the AWH
466 region may or may not follow externally forced changes: (1) A long-term warming trend differ-
467 ence, (2) a different spectrum of high-frequency SST variability, or (3) true internal low-frequency
468 variability. To investigate whether (3) is a viable explanation, we applied running means from 5-20
469 years (Fig. 8e, h, m; middle panel), we linearly detrended model and observations (Fig. 8c), or we
470 did both (Fig. 8f, j, n; rhs panel).

471 What we find is that the AWH is robust against temporal averaging as far as non-detrended
472 data are concerned. In contrast, if detrended data are used, the temporal averaging aligns the
473 NAVI/AWH region with the NHem forcing response in terms of correlation, maintaining its
474 (forced) multidecadal low-frequency variability. In fact, detrending alone considerably reduces
475 the unique behaviour of the AWH region already (Fig. 8c). What we infer from this is that the
476 secular warming trend (1) is responsible for the specific characteristic of the AMV region. The
477 root cause of this cooling trend is well known and one of the key features in GCM projections
478 (Rahmstorf et al. 2015; Menary and Wood 2018). The high-frequency variability over the wider

NA region is higher than on global average, but comparable in magnitude to the western North Pacific (equally high supply of baroclinicity) or Eurasia (Fig. 8b).

After a decade, not much multi-annual stochastic variability is left (Fig. 8e). Together with the Indian ocean, the wider NA region shows high correlations with the NHem model after trend removal in both (Fig. 8f), suggesting substantial dependencies on externally forced low-frequency variability. It is a different story over land (and much of the Pacific), where the signal-to-noise ratio is lower on decadal scales due to limited radiative constraints on winter temperatures (Cohen et al. 2012; Knutson et al. 2013; Deser et al. 2017). The positive correlation between the 20 year low-pass filtered, detrended AMV and the Arctic (Fig. 8k) is physically very plausible as amplified Arctic warming relies on heat transport via the NA region, governed by the NAO index and the associated strength of the AMOC. However, it is the forced long-term warming trend that is the driver as evident from Figs. 8m and 8n.

Since no noticeable low-frequency signal can be detected over the key AWH region (Fig. 8n), we conclude that it is unlikely that internal variability on timescales > 5 years plays an important role in the North Atlantic. There is room for 1-5 year unforced feedbacks, but apart from the cooling due to the long-term decline in AMOC strength (Fig. 8m), high- and low-frequency AMV pattern appear to be externally forced according to our response model results. This is in line with an empirical model study that uses multiple regression to attribute forcing contributions globally (Suckling et al. 2017), and also supported by other studies that show that subpolar NA variability is largely driven by AMOC changes, with little evidence for a strong AMV-AMOC link (Marini and Frankignoul 2014; Frankignoul et al. 2017).

In conclusion, combined with the recent downward trend in the new NAVI index, our analysis strongly suggests that the impact of internally generated NA ocean dynamics on Global, NHem and Land temperatures is rather limited. Remaining AMOC related to low-frequency variability

(Zhang 2017) may have regional implications, but a strong influence beyond that is unlikely. The results are supported by another simple exercise in which NA SSTs are weighted by the surface area of the AMV/NAVI region, divided by the NHem surface area. This way, the fractional fingerprint of the AMV on NHem temperatures can be inferred. The peak contribution would be $<0.03^{\circ}\text{C}$, assuming all NA SST variability is of internal origin, which we have shown not to be a very plausible conjecture. Helped by a more advanced (yet still debatable) regression analysis, we note that Folland et al. (2018) also found almost no AMV contribution to global temperature.

7. Conclusions

With explained variabilities of observed global temperatures of up to 98% (30 year smooth) or $\sim 93\%$ (with ENSO variability), respectively, our impulse response model performs exceptionally well. We are able to match the historical temperature evolution since at least 1850 in general, and succeed in reproducing both the EW and the MC period with high precision in particular, without the need to invoke unexplained internal multidecadal temperature variability as an additional driver.

Three key aspects are crucial for an appropriate attribution of the temperature response to external radiative forcing perturbations. (1) Careful treatment of the spatially heterogeneous AER forcing as its temporal evolution has major repercussions for both the EW and the MC period. (2) Removal of the WWII warm bias in the current generation of SST datasets as there is now solid evidence that 1942-1945 period is biased warm to differing degrees, causing a spurious warming trend at the end of the EW period. (3) Calibration of the fast response time in order to account for the mediating effects of ENSO as far as the response to volcanic eruptions is concerned.

While others (Mann et al. 2014; Folland et al. 2018) have found similarly good agreement as far as the GMST evolution is concerned, our analysis demonstrates that it is possible to reproduce

the temperature evolution separately for NHem, SHem, Land and Ocean with equal precision. We achieve this by introducing a set of suitable TCR calibration factors that are informed by observed (HadOST) and modelled (HadCM3) TWRs and traceable throughout the analysis. Apart from minor fine-tuning related to the deduced TWR for AER, every response model parameter used in our study is backed up by independent analysis and/or based on well-established research. The use of updated aerosol emission and volcanic forcing data as well as the application of a longer fast response time (complemented by a hemispherically more uniform fast VOL response) are otherwise the only changes that we made compared to previous iterations within the response model framework. Owing to the introduced analytical constraints, which are designed to avoid model tuning, our results warrant robustness against overfitting.

With the introduction of HadOST, which includes a coastal temperature analysis inspired by Cowtan et al. (2017) that appears least biased with regard to the incorporated HadISST2 and OSTIA SSTs, we add another option to the existing batch of GMST datasets. We recommend to use it more widely as it resolves some of the discrepancies present in HadSST3 before 1940. Despite a smaller warm bias during WWII in HadISST2 compared to HadSST3, we still have to impose a correction factor (-0.08°C for GMST) to reconcile it with the coastal hybrid temperature time-series. As a result, almost all of the EW warming could ultimately explained by external forcing changes, which - if confirmed by future research - may call the current partition of attributable EW causes, as recently reviewed in Hegerl et al. (2018), in considerable doubt.

In our assessment of potential contributions from Atlantic and Pacific multidecadal variability, we demonstrate that with the exception of prolonged periods of El Niño or La Niña preponderance, there is little room for internal unforced ocean variability beyond subdecadal timescales, which is particularly true for the NA region. This finding is buttressed by our demonstration that despite high co-variability, cause (VOL and AER) and effect (AMV) are clearly distinguishable. That

550 does not mean AMV cannot have internal mechanisms (Zhang 2017), rather only that the signal
551 cannot be detected in Global or NHem (nor is necessary to explain their temporal evolution).
552 Hence the traditional AMV index must not be used as predictor or explanatory variable, as it may
553 lead to demonstrably incorrect or flawed attribution results (Hetzinger et al. 2008; Chylek et al.
554 2009; Huss et al. 2010; Wyatt et al. 2012; Tung and Zhou 2013; Chylek et al. 2014; Pasini et al.
555 2016; Hodgkins and Wilson 2017; Yan et al. 2017; Shen et al. 2018; Zhang et al. 2018). We
556 suggest a revised AMV index formulation (NAVI) which avoids such pitfalls as it better mirrors
557 the long-term AMOC decline as suggested, for example, in (Rahmstorf et al. 2015).

558 On that note, we also caution against confusing atmospherically driven short-term variability
559 (noise) with changes due to anthropogenic or natural external forcing factors (signal). As demon-
560 strated in the supplementary analysis, anomalous atmospheric NHem winter circulation features
561 explain most of the short-term AMOC variability, acting as the control knob on multi-monthly
562 timescales. Longer timescales are conceivable: (1) Via changing wind stress related to anoma-
563 lous NAO phasing, which in turn affects the subpolar horizontal gyre circulation (Piecuch et al.
564 2017). (2) Via atmospheric teleconnections associated with ENSO such as the PNA-NAO rela-
565 tionship (Pinto et al. 2011), which in turn links to the emerging paradigm of the Pacific basin as
566 pacemaker for global temperature (Guan and Nigam 2009; Kosaka and Xie 2013; England et al.
567 2014; Schurer et al. 2015; Dong and McPhaden 2017; Frankignoul et al. 2017). (3) Via *ocean*
568 *memory* effects, which may favour the reoccurrence of certain large-scale weather patterns in the
569 Euro-Atlantic region during successive boreal winter seasons via air-sea coupling (Scaife et al.
570 2014). But generally, progress in understanding Atlantic decadal climate variability has been slow
571 (Yeager and Robson 2017). Taken together, our analysis underscores that despite the complexities
572 of the climate system, changes to the mean state are dominated by radiative forcings on longer
573 timescales and ENSO-related variability on shorter timescales.

574 By virtue of these findings, we are confident that our associated best TCR estimate of 1.57
575 (± 0.70) K is robust, despite a substantial error range due to the large forcing uncertainty. We
576 strongly advise against the use of ECS estimates based on the instrumental record alone with-
577 out considering further evidence (from paleo-archives or GCMs), as they cannot be reliably con-
578 strained with data of such a short time interval.

579 In a future analysis, we aim to quantify another important response model feature which also
580 contributes to an improved representation of the EW period. In a nutshell, it can be demonstrated
581 that failure to initialise the response model (or GCMs for that matter) before a series of strong
582 volcanic eruptions will very likely bias the beginning of the simulated EW period warm, leading
583 to an artificially low warming trend in models.

584 *Acknowledgments.* The authors thank the UK Met Office Hadley Centre (Had-
585 CRUT4, HadSST3, HadISST2), the National Centre for Ocean Forecasting (OSTIA), NASA/GISS
586 (GISTEMP), and Berkeley Earth (BE) for providing the observed temperature and SST data.
587 We also thank the National Oceanic and Atmospheric Administration (NOAA) for providing the
588 surface wind vector, 850hPa temperature and Optimum Interpolation (OI) SST data, as well as the
589 NOAA Earth System Research Laboratory (ESRL) for providing the 20th Century reanalysis data.
590 Furthermore, we thanks Piers Forster for providing the updated radiative forcing data, including
591 the forcing uncertainty ensemble, and the Greenwich Royal Observatory for providing the sunspot
592 number data. Finally, we would like to acknowledge the input from M.R. Allen regarding the
593 original conceptual idea in context of the Global Warming Index, as well as the input from B.B.B.
594 Both, who provided valuable comments during the early stages of paper writing. This work was
595 supported and partly funded by the UK National Environmental Research Council's DOCILE
596 project under the grant assignment NE/P002099/1 (2019-1552).

Calculation of coupling factors

As outlined in Section 3, inter-hemispheric energy exchanges in response to the heterogeneous distribution of AER need to be balanced by virtue of so-called coupling factors. As shown in Fig. 3, for AER there is a notable discrepancy between the TCR scaling factors (2.5/1.9 for NHem/SHem and 2.7/1.95 for Land/Ocean) and the diagnosed AER-TWR (5.1 for NHem/SHem and 2.9 for Land/Ocean). A discrepancy that does not appear for WMGHG and VOL (for instance: $TWRD_{WMGHG} = 2.23/0.97 = 2.3$). This is where the coupling factor comes into play. Since the TCR scaling only works under the assumption that the NHem aerosol response is governed by NHem aerosol emissions (and vice versa for SHem; same problem for Land and Ocean), we have to find a way to accommodate the additional temperature response from SHem emissions in case of the NHem response.

We have applied two-stage methodology: (1) We derive the emission ratio, i.e. the ratio between the hemispheric (and land/ocean) and the total global aerosol emission strength. It is a function of the fractional contribution of each hemisphere (or land/ocean) and could vary between 95% to 5%, up until 50% to 50%. (2) We determine the optimal fractional contribution or coupling strength. For this to work, we balance the ratio of the TCR scaling factors (e.g. 2.5 for NHem and 1.9 for SHem) and the TWRD (e.g. 5.1 for NHem/SHem). Since the effective forcing of the SHem emissions will be lower (due to the lower TCR scaling factor), a secondary scaling factor has to be applied to NHem and SHem emission strength. This factor is only be equal for one particular set of coupling strengths. For NHem/SHem, the fractional contributions are 85% and 15%, associated with a NHem emission ratio (fractional NHem emission divided by Global emissions) of 1.58 and a SHem emission ratio (fractional SHem emission divided by Global emissions) of 0.42. The

620 additional secondary scaling factor is 0.93. Hence the NHem/SHem emission ratio is reduced to
621 1.47 and 0.38. Their ratio defines the coupling factor, which is 3.9 accordingly ($=1.47/0.38$).

622 Applying the same method for Land/Ocean, the fractional contributions are 70% and 30%, asso-
623 ciated with a Land emission ratio (fractional Land emission divided by Global emissions) of 1.35
624 and an Ocean emission ratio (fractional Ocean emission divided by Global emissions) of 0.65. The
625 secondary scaling factor is 1.08, which is explained by the fact that the Ocean sensitivity is lower,
626 but instead covers a much larger area fraction compared to Land (area fractions are equal in case
627 of NHem/SHem). Hence the Land/Ocean emission is increased to 1.46 and 0.70. The associated
628 ratio to determine the coupling factor is 2.1 (see lower box in Fig. 3 and fine-print below it).

629 APPENDIX B

630 **Decadal temperature evolution**

631 As highlighted in the main text, SST observations are still afflicted with considerable uncertain-
632 ties. Having investigated time series of field means, here we provide the spatiotemporal context
633 and discuss potential causes for some of the discrepancies noted above. In Fig. B1, the GMST
634 dataset used in this study are plotted as decadal average from 1850-1859 to 2010-2017 (BE,
635 Cru4CW, HadOST), accompanied by GISTEMP during 1880-1889 to 2010-2017. In addition,
636 the 20th Century reanalysis (20C Rean) (Hirahara et al. 2014), the ensemble mean of a subset of
637 CMIP5 simulations are plotted, together with NorESM1-M (Bentsen et al. 2013) which is found
638 to represent the temperature evolution since 1850 very well compared to observations. We also
639 added the recently proposed Hybrid SSTs (Cowtan et al. 2017). The decade 1940-1949 is high-
640 lighted by the red box. Note that HadOST and Cru4CW use the same infilled HadCRUT4 (Morice
641 et al. 2012) data over land. We also note that both, Cru4CW and BE have shown to carry a small

negative trend bias in recent years (Hausfather et al. 2017). If that is not enough, it has also been suggested that the Arctic region might still be biased cold (Wang et al. 2017; Way et al. 2017).

As noted above, the 1880-1935 period is too cold in HadSST3 (Cru4CW) compared to HadISST2 (HadOST), most pronounced during 1890-1920. Looking at those three decades, at least the first two show a noteworthy feature near Cape Cauldron off the southern tip of South Africa, presumably associated with the Agulhas and Brazil currents. The otherwise distinct cold SST anomalies in the turbulent exit region where the Agulhas current leaks into the South Atlantic ocean (compare HadOST) turns into a vast area of cold SST anomalies that essentially covers most of the South Atlantic. Given the poor observational coverage and the intrinsic shortcomings of any infilling technique (Kriging in case of Cru4CW), it is likely that the cold South Atlantic SST anomalies in Cru4CW, BE and GISTEMP are exaggerated to varying degrees.

Since the only bias in HadOST with regard to our response model results was found during the 1850-1879 period, mainly caused by warmer NH SST conditions, it is interesting to ask whether the warm SST anomaly in the North Pacific in HadOST is real given it does not show up in other observational datasets. While such a pattern is consistent with a prolonged PDV negative phase, the amplitude of the anomaly appears very strong, especially during the 1860s. The pattern re-occurs during the 1950-80 period, but background SSTs are less cold than during the 1850-79 period. This is arguably a feature which deserves to be investigated in more detail, particularly in light of recent work by Huang et al. (2018).

Regarding the WWII period, even though we have plotted decadal averages, what stands out is the sudden warming of all ocean basins in ERSST during the 1940s (and to a much lesser extent HadSST3 and HadOST). As evident from Fig. 5a, Land did not notably warm during the same period, which strongly suggests an artifactual feature, related to biases due to the previously explained change in fleet composition during WWII. The final feature we would like to mention

666 concerns the cold bias during 1950-80 in ERSST. While the general NHem ocean cooling due
667 to increased anthropogenic SO₂ emissions is visible in all observations (and CMIP5 simulations,
668 irrespective of some temporal misalignments), ERSST seems to exaggerate the cooling slightly
669 given that the spatial pattern of the SST anomalies are indistinguishable from other datasets. We
670 speculate that this might be a general theme in ERSST given that it draws heavily from maritime
671 nighttime measurements in contrast to other products. We note that ERSSTv5 (Huang et al. 2017)
672 has not changed notably compared to ERSSTv4.

673 As we cannot provide robust conclusions with regard to causes for the mismatch between dif-
674 ferent observational dataset at this point, we would like to close by encouraging the research
675 community to address the oftentimes under-appreciated problems in more depth. Insights from
676 energy balance models such as presented in this study can guide such efforts. With clever new
677 strategies to combine the various information available, we are confident that the remaining gaps
678 in our understanding will be eliminated and instrumental data brought into better agreement.

679 **References**

- 680 Aldrin, M., M. Holden, P. Guttorp, R. Skeie, G. Myhre, and T. Berntsen, 2012: Bayesian estima-
681 tion of climate sensitivity based on a simple climate model fitted to observations of hemispheric
682 temperatures and global ocean heat content. *Environmetrics*, **23** (3), 253–271.
- 683 Andrews, T., and M. Webb, 2018: The Dependence of Global Cloud and Lapse Rate Feedbacks
684 on the Spatial Structure of Tropical Pacific Warming. *Journal of Climate*, **31**, 641–654.
- 685 Arfeuille, F., D. Weisenstein, H. Mack, E. Rozanov, T. Peter, and S. Brönnimann, 2014: Volcanic
686 forcing for climate modeling: a new microphysics-based data set covering years 1600–present.
687 *Climate of the Past*, **10**, 359–375, doi:10.5194/cp-10-359-2014.

688 Armour, K., 2017: Energy budget constraints on climate sensitivity in light of inconstant climate
689 feedbacks. *Nature Climate Change*, **7**, 331–335.

690 Armour, K., J. Marshall, J. Scott, A. Donohoe, and E. Newsom, 2016: Southern Ocean warming
691 delayed by circumpolar upwelling and equatorward transport. *Nature Geoscience*, **9**, 549–554.

692 Bellomo, K., L. Murphy, M. Cane, A. Clement, and L. Polvani, 2018: Historical forcings as main
693 drivers of the Atlantic multidecadal variability in the CESM large ensemble. *Climate Dynamics*,
694 **50 (9-10)**, 3687–3698.

695 Bellucci, A., A. Mariotti, and S. Gualdi, 2017: The role of forcings in the 20th century North
696 Atlantic multi-decadal variability: the 1940-1975 North Atlantic cooling case study. *Journal of*
697 *Climate*.

698 Bentsen, M., and Coauthors, 2013: The Norwegian Earth System Model, NorESM1-M - Part 1:
699 Description and basic evaluation of the physical climate. *Geoscientific Model Development*, **6**,
700 687–720.

701 Bindoff, N. L., and Coauthors, 2013: *Detection and Attribution of Climate Change: from Global*
702 *to Regional. In: Climate Change 2013: The Physical Science Basis. Contribution of Working*
703 *Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.*
704 Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

705 Boers, R., T. Brandsma, and A. Pier Siebesma, 2017: Impact of aerosols and clouds on decadal
706 trends in all-sky solar radiation over the Netherlands (1966-2015). *Atmospheric Chemistry and*
707 *Physics*, **17**, 8081–8100.

708 Bond, T., and Coauthors, 2013: Bounding the role of black carbon in the climate system: A
709 scientific assessment. *Journal of Geophysical Research: Atmospheres*, **118 (11)**, 5380–5552.

Booth, B., N. Dunstone, P. Paul R. Halloran, T. Andrews, and N. Bellouin, 2012: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability. *Nature*, **484**, 228–232.

Boucher, O., and M. Reddy, 2008: Climate trade-off between black carbon and carbon dioxide emissions. *Energy Policy*, **36**, 193–200.

Brönnimann, S., 2009: Early twentieth-century warming. *Nature Geoscience Correspondance*, **2**, 735–736.

Cai, W., and Coauthors, 2015: ENSO and greenhouse warming. *Nature Climate Change*, **5**, 849–859.

Cane, M., A. Clement, L. Murphy, and K. Bellomo, 2017: Low-Pass Filtering, Heat Flux, and Atlantic Multidecadal Variability. *Journal of Climate*, **30**, 7529–7553.

Carslaw, K., and Coauthors, 2013: Large contribution of natural aerosols to uncertainty in indirect forcing. *Nature*, **503 (7474)**, 67–71.

Cheng, L., K. Trenberth, M. Palmer, J. Zhu, and J. Abraham, 2016: Observed and simulated full-depth ocean heat-content changes for 1970-2005. *Ocean Science*, **12**, 925–935.

Chung, E.-S., and B. Soden, 2017: Hemispheric climate shifts driven by anthropogenic aerosol-cloud interactions. *Nature Geoscience*, **10**, 566–572, doi:10.1038/ngeo2988.

Chylek, P., C. K. Folland, G. Lesins, M. K. Dubey, and M. Wang, 2009: Arctic air temperature change amplification and the Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, **36 (14)**, L14 801.

730 Chylek, P., J. Klett, G. Lesins, M. Dubey, and N. Hengartner, 2014: The Atlantic Multidecadal
731 Oscillation as a dominant factor of oceanic influence on climate. *Geophysical Research Letters*,
732 **41 (5)**, 1689–1697.

733 Clement, A., K. Bellomo, L. Murphy, M. Cane, T. Mauritsen, G. Rädel, and B. Stevens, 2015:
734 The Atlantic Multidecadal Oscillation without a role for ocean circulation. *Science*, **350 (6258)**,
735 320–324.

736 Clement, A., M. Cane, L. Murphy, K. Bellomo, T. Mauritsen, and B. Stevens, 2016: Response
737 to Comment on "The Atlantic Multidecadal Oscillation without a role for ocean circulation".
738 *Science*, **352 (6293)**, 1527b.

739 Cohen, J. L., J. C. Furtado, M. Barlow, V. A. Alexeev, and J. E. Cherry, 2012: Asymmetric seasonal
740 temperature trends. *Geophysical Research Letters*, **39 (4)**, L04 705.

741 Cohen, J. L., and Coauthors, 2014: Recent Arctic amplification and extreme mid-latitude weather.
742 *Nature Geoscience*, **7**, 627–637.

743 Cowtan, K., R. Rohde, and Z. Hausfather, 2017: Estimating biases in Sea Surface Temperature
744 records using coastal weather stations. *Quarterly Journal of the Royal Meteorological Society*,
745 **in review**.

746 Cowtan, K., and R. Way, 2014: Coverage bias in the HadCRUT4 temperature series and its impact
747 on recent temperature trends. *Quarterly Journal of the Royal Meteorological Society*, **140 (683)**,
748 1935–1944.

749 Cowtan, K., and Coauthors, 2015: Robust comparison of climate models with observations using
750 blended land air and ocean sea surface temperatures. *Geophysical Research Letters*, **42 (15)**,
751 6526–6534.

752 Crowley, T., and M. Unterman, 2013: Technical details concerning development of a 1200
 753 yr proxy index for global volcanism. *Earth System Science Data*, **5**, 187–197, doi:10.5194/
 754 essd-5-187-2013.

755 Crowley, T., G. Zielinski, B. Vinther, R. Udisti, K. Kreutz, J. Cole-Dai, and E. Castellano, 2008:
 756 Volcanism and the little ice age. *PAGES News*, **16 (2)**, 22–23.

757 Dallafior, T., D. Folini, R. Knutti, and M. Wild, 2016: Mixed-layer ocean responses to anthro-
 758 pogenic aerosol dimming from 1870 to 2000. *Journal of Geophysical Research: Atmospheres*,
 759 **121**, 49–66, doi:10.1002/2015JD024070.

760 Davis, L.L.B., D.W.J. Thompson, J.J.. Kennedy, and E.C. Kent, 2019: The importance of unre-
 761 solved biases in 20th century sea-surface temperature observations. *Bulletin of the American*
 762 *Meteorological Society*, **99 (12)**, doi:10.1175/BAMS-D-18-0104.1.

763 de Boissésón, E., M. Balmaseda, S. Abdalla, E. Källén, and P. Janssen, 2014: How robust is the
 764 recent strengthening of the Tropical Pacific trade winds? *Geophysical Research Letters*, **41**,
 765 4398–4405.

766 Delworth, T., and M. Mann, 2000: Observed and simulated multidecadal variability in the North-
 767 ern Hemisphere. *Climate Dynamics*, **16**, 661–676.

768 Deser, C., R. Guo, and F. Lehner, 2017: The relative contributions of tropical Pacific sea surface
 769 temperatures and atmospheric internal variability to the recent global warming hiatus. *Geophys-
 770 ical Research Letters*, **44 (15)**, 7945–7954.

771 Deser, C., and Coauthors, 2012: ENSO and pacific decadal variability in the community climate
 772 system model version 4. *Journal of Climate*, **25 (8)**, 2622–2651.

- 773 Dewitte, S., and S. Nevens, 2016: The total solar irradiance climate data record. *The Astrophysical*
774 *Journal*, **830** (25), 1–5.
- 775 DiNezio, P., B. Kirtman, A. Clement, S.-K. Lee, G. Vecchi, and A. Wittenberg, 2012: Mean
776 Climate Controls on the Simulated Response of ENSO to Increasing Greenhouse Gases. *Journal*
777 *of Climate*, **25**, 7399–7420.
- 778 Ding, Y., J. Carton, G. Chepurin, G. Stenchikov, A. Robock, L. Sentman, and J. Krasting, 2014:
779 Ocean response to volcanic eruptions in Coupled Model Intercomparison Project 5 simulations.
780 *Journal of Geophysical Research: Oceans*, **119**, 1–16.
- 781 Dong, L., and M. McPhaden, 2017: The role of external forcing and internal variability in regulat-
782 ing global mean surface temperatures on decadal timescales. *Environmental Research Letters*,
783 **12**, 034 011.
- 784 Dong, L., T. Zhou, and X. Chen, 2014: Changes of Pacific decadal variability in the twentieth
785 century driven by internal variability, greenhouse gases, and aerosols. *Geophysical Research*
786 *Letters*, **41**, 8570–8577.
- 787 Donlon, C. J., M. Martin, J. Stark, J. Roberts-Jones, E. Fiedler, and W. Wimmer, 2012: The
788 Operational Sea Surface Temperature and Sea Ice Analysis (OSTIA) system. *Remote Sensing*
789 *of the Environment*, **116**, 140–158, doi:10.1016/j.rse.2010.10.017.
- 790 Ducheze, A., and Coauthors, 2016: Drivers of exceptionally cold North Atlantic Ocean temper-
791 atures and their link to the 2015 European heat wave. *Environmental Research Letters*, **11**,
792 074 004.

793 Dumitrescu, A., C. Gueymard, and V. Badescu, 2017: Reconstruction of historical aerosol optical
794 depth time series over Romania during summertime. *International Journal of Climatology*, **37**,
795 in press.

796 Emile-Geay, J., R. Seager, M. Cane, E. Cook, and G. Haug, 2008: Volcanoes and ENSO over the
797 Past Millennium. *Journal of Climate*, **21**, 3134–3148.

798 England, M., and Coauthors, 2014: Recent intensification of wind-driven circulation in the Pacific
799 and the ongoing warming hiatus. *Nature Climate Change*, **4**, 222–227.

800 Etminan, M., G. Myhre, E. Highwood, and K. Shine, 2016: Radiative forcing of carbon dioxide,
801 methane, and nitrous oxide: A significant revision of the methane radiative forcing. *Geophysical*
802 *Research Letters*, **43**, 12,614–12,623.

803 Flato, G., and Coauthors, 2013: *Evaluation of Climate Models*. In: *Climate Change 2013: The*
804 *Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the*
805 *Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United
806 Kingdom and New York, NY, USA.

807 Folland, C., O. Boucher, and D. Colman, A. and Parker, 2018: Causes of irregularities in trends of
808 global mean surface temperature since the late 19th century. *Science Advances*, **4** (6), eaao5297.

809 Forest, C., 2018: Inferred Net Aerosol Forcing Based on Historical Climate Changes: a Review.
810 *Current Climate Change Reports*, **4** (1), 11–22, doi:10.1007/s40641-018-0085-2.

811 Forster, P., T. Andrews, P. Good, J. Gregory, L. Jackson, and M. Zelinka, 2013: Evaluating ad-
812 justed forcing and model spread for historical and future scenarios in the CMIP5 generation of
813 climate models. *Journal of Geophysical Research*, **118**, 1–12.

814 Foster, G., and S. Rahmstorf, 2011: Global temperature evolution 1979-2010. *Environmental Re-*
815 *search Letters*, **6**, 044 022.

816 Francis, J., 2017: Why Are Arctic Linkages to Extreme Weather Still Up in the Air? *Bulletin of*
817 *the American Meteorological Society*, **98**, in press.

818 Frankignoul, C., G. Gastineau, and Y.-O. Kwon, 2017: Estimation of the SST response to anthro-
819 pogenic and external forcing, and its impact on the Atlantic Multidecadal Oscillation and the
820 Pacific Decadal Oscillation. *Journal of Climate*, **30**, in press.

821 Geoffroy, O., D. Saint-Martin, G. Bellon, A. Voldoire, D. Olivie, and S. Tyteca, 2013a: Transient
822 Climate Response in a Two-Layer Energy-Balance Model. Part II: Representation of the Effi-
823 cacy of Deep-Ocean Heat Uptake and Validation for CMIP5 AOGCMs. *Journal of Climate*, **26**,
824 1859–1876.

825 Geoffroy, O., D. Saint-Martin, D. Olivie, A. Voldoire, G. Bellon, and S. Tyteca, 2013b: Transient
826 Climate Response in a Two-Layer Energy-Balance Model. Part I: Analytical Solution and Pa-
827 rameter Calibration Using CMIP5 AOGCM Experiments. *Journal of Climate*, **26**, 1841–1857.

828 Gregory, J., T. Andrews, and P. Good, 2015: The inconstancy of the transient climate response
829 parameter under increasing CO₂. *Philosophical Transactions of the Royal Society A*, **373 (2054)**,
830 20140 417.

831 Guan, B., and S. Nigam, 2009: Analysis of Atlantic SST Variability Factoring Interbasin Links
832 and the Secular Trend: Clarified Structure of the Atlantic Multidecadal Oscillation. *Journal of*
833 *Climate*, **22**, 4228–4240.

834 Hansen, J., R. Ruedy, M. Sato, and K. Lo, 2010: Global surface temperature change. *Reviews of*
835 *Geophysics*, **48 (RG4004)**, 1–29.

836 Hansen, J., and Coauthors, 2016: Ice melt, sea level rise and superstorms: evidence from pale-
837 oclimate data, climate modeling, and modern observations that 2C global warming could be
838 dangerous. *Atmospheric Chemistry and Physics*, **16**, 3761–3812.

839 Hausfather, Z., K. Cowtan, D. Clarke, P. Jacobs, M. Richardson, and R. Rohde, 2017: Assessing
840 recent warming using instrumentally homogeneous sea surface temperature records. *Science*
841 *Advances*, **3** (1), e1601 207.

842 Haustein, K., M. Allen, P. Forster, F. Otto, D. Mitchell, H. D. Matthews, and D. Frame, 2017: A
843 robust real-time Global Warming Index. *Nature Scientific Reports*, **7** (15417), 1–6.

844 Hay, S., P. Kushner, R. Blackport, and K. McCusker, 2018: On the relative robustness of the
845 climate response to high-latitude and low-latitude warming. *Geophysical Research Letters*, **ac-**
846 **cepted**.

847 Hegerl, G., S. Brönnimann, A. Schurer, and T. Cowan, 2018: The early 20th century warming:
848 Anomalies, causes, and consequences. *WIREs Climate Change*, **2018**;e522, 1–20, doi:10.1002/
849 wcc.522.

850 Henley, B., 2017: Pacific decadal climate variability: Indices, patterns and tropical-extratropical
851 interactions. *Global and Planetary Change*, **155**, 42–55.

852 Hetzinger, S., M. Pfeiffer, W.-C. Dullo, N. Keenlyside, M. Latif, and J. Zinke, 2008: Caribbean
853 coral tracks Atlantic Multidecadal Oscillation and past hurricane activity. *Geology*, **36** (1), 11–
854 14.

855 Hirahara, S., M. Ishii, and Y. Fukuda, 2014: Centennial-Scale Sea Surface Temperature Analysis
856 and Its Uncertainty. *Journal of Climate*, **27**, 57–75.

857 Hodgkins, G., and D. Wilson, 2017: Climate-driven variability in the occurrence of major floods
858 across North America and Europe. *Journal of Hydrology*, **552**, 704–717.

859 Hoesly, R., and Coauthors, 2017: Historical (1750-2014) anthropogenic emissions of reactive
860 gases and aerosols from the Community Emission Data System (CEDS). *Geoscientific Model
861 Development*, **11**, 369–408.

862 Huang, B., W. Angel, and T. Boyer, 2018: Evaluating SST Analyses with Independent Ocean
863 Profile Observations. *Journal of Climate*, doi:10.1175/JCLI-D-17-0824.1.

864 Huang, B., and Coauthors, 2017: Extended Reconstructed Sea Surface Temperature version 5
865 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate*, **30** (21), in press,
866 doi:10.1175/JCLI-D-16-0836.1.

867 Huber, M., and R. Knutti, 2014: Natural variability, radiative forcing and climate response in the
868 recent hiatus reconciled. *Nature Geoscience*, **7**, 651–656.

869 Hurrell, J. W., and C. Deser, 2009: North Atlantic climate variability: the role of the North Atlantic
870 Oscillation. *Journal of Maritime Systems*, **78**, 28–41.

871 Huss, M., R. Hock, A. Bauder, and M. Funk, 2010: 100-year mass changes in the Swiss Alps
872 linked to the Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, **37** (10), L10 501.

873 Iwi, A., L. Hermanson, K. Haines, and R. Sutton, 2012: Mechanisms Linking Volcanic Aerosols
874 to the Atlantic Meridional Overturning Circulation. *Journal of Climate*, **25**, 3039–3051.

875 Johannessen, O. M., and Coauthors, 2004: Arctic climate change: observed and modelled temper-
876 ature and sea-ice variability. *Tellus A*, **56**, 328–341.

877 Jones, G., P. Stott, and N. Christidis, 2013: Attribution of observed historical near-surface tem-
 878 perature variations to anthropogenic and natural causes using CMIP5 simulations. *Journal of*
 879 *Geophysical Research: Atmospheres*, **118**, 4001–4024.

880 Jones, G., P. Stott, and J. Mitchell, 2016: Uncertainties in the attribution of greenhouse gas warm-
 881 ing and implications for climate prediction. *Journal of Geophysical Research: Atmospheres*,
 882 **121 (12)**, 6969–6992.

883 Josey, S.A., M.F. de Jong, M. Oltmanns, G.K. Moore, and R.A. Weller, 2019: Extreme Variability
 884 in Irminger Sea Winter Heat Loss Revealed by Ocean Observatories Initiative Mooring and the
 885 ERA5 Reanalysis. *Geophysical Research Letters*, **46**, 1–10.

886 Kajtar, J., A. Santoso, S. McGregor, M. England, and Z. Baillie, 2017: Model under-representation
 887 of decadal Pacific trade wind trends and its link to tropical Atlantic bias. *Climate Dynamics*, **49**,
 888 1–14.

889 Karl, T., and Coauthors, 2015: Possible artifacts of data biases in the recent global surface warming
 890 hiatus. *Science*, **348 (6242)**, 1469–1472.

891 Kennedy, J., N. Rayner, S. Millington, and M. Saunby, 2017: The Met Office Hadley Centre sea
 892 ice and sea surface temperature data set, version 2: 2. Sea surface temperature analysis. *Journal*
 893 *of Geophysical Research: Atmospheres*, **in preparation**.

894 Kennedy, J., N. Rayner, R. Smith, M. Saunby, and P. D.E., 2011: Reassessing biases and other
 895 uncertainties in sea-surface temperature observations measured in situ since 1850: 2. Biases and
 896 homogenization. *Journal of Geophysical Research*, **116 (D14)**, D14 104.

897 Kent, E., and Coauthors, 2017: A Call for New Approaches to Quantifying Biases in Observations
 898 of Sea Surface Temperature. *Bulletin of the American Meteorological Society*, **98 (9)**, in press.

899 Klimont, Z., S. Smith, and J. Cofala, 2013: The last decade of global anthropogenic sulfur dioxide:
900 2000-2011 emissions. *Environmental Research Letters*, **8**, 014 003.

901 Knight, J., R. Allan, C. Folland, M. Vellinga, and M. Mann, 2005: A signature of persistent natural
902 thermohaline circulation cycles in observed climate. *Geophysical Research Letters*, **32 (20)**,
903 L20 708.

904 Knight, J., C. Folland, and A. Scaife, 2006: Climate impacts of the Atlantic Multidecadal Oscilla-
905 tion. *Geophysical Research Letters*, **33 (17)**, L17 706.

906 Knudsen, M., B. Jacobsen, M.-S. Seidenkrantz, and J. Olsen, 2014: Evidence for external forcing
907 of the Atlantic Multidecadal Oscillation since termination of the Little Ice Age. *Nature Com-*
908 *munications*, **5 (3323)**.

909 Knutson, T., F. Zeng, and A. Wittenberg, 2013: Multimodel Assessment of Regional Surface
910 Temperature Trends: CMIP3 and CMIP5 Twentieth-Century Simulations. *Journal of Climate*,
911 **26**, 8709–8743.

912 Knutti, R., M. Rugenstein, and G. Hegerl, 2017: Beyond equilibrium climate sensitivity. *Nature*
913 *Geoscience*, **10**, 727–736.

914 Kohyama, T., and D. Hartmann, 2017: No Access Nonlinear ENSO Warming Suppression
915 (NEWS). *Journal of Climate*, **30**, 4227–4251.

916 Kosaka, Y., and S.-P. Xie, 2013: Recent global-warming hiatus tied to equatorial Pacific surface
917 cooling. *Nature*, **501**, 403–407.

918 Kunkel, D., H. Tost, and M. Lawrence, 2013: Aerosol pollution potential from major population
919 centers. *Atmospheric Chemistry and Physics*, **13**, 4203–4222.

- 920 Le, T., 2017: ENSO response to external forcing in CMIP5 simulations of the last millennium.
921 *Global and Planetary Change*, **148**, 105–112.
- 922 Lean, J., 2018: Observation-based detection and attribution of 21st century climate change. *WIREs*
923 *Climate Change*, **9:e511**, 1–20.
- 924 Lean, J., and D. Rind, 2008: How natural and anthropogenic influences alter global and regional
925 surface temperatures: 1889 to 2006. *Geophysical Research Letters*, **35 (L18701)**.
- 926 Levine, A., D. Frierson, and M. McPhaden, 2018: AMO forcing of Multidecadal Pacific ITCZ
927 variability. *Journal of Climate*, **in press**, doi:10.1175/JCLI-D-17-0810.1.
- 928 Lewis, N., and J. Curry, 2015: The implications for climate sensitivity of AR5 forcing and heat
929 uptake estimates. *Climate Dynamics*, **45 (3-4)**, 1009–1023.
- 930 Lewis, N., and J. Curry, 2018: The impact of recent forcing and ocean heat uptake data on esti-
931 mates of climate sensitivity. *Journal of Climate*, **in press**, doi:10.1175/JCLI-D17-0667.1.
- 932 L’Heureux, M. L., S. Lee, and B. Lyon, 2013: Recent multidecadal strengthening of the Walker
933 circulation across the tropical Pacific. *Nature Climate Change*, **3**, 571–576.
- 934 Li, S., and A. Jarvis, 2009: Long run surface temperature dynamics of an AOGCM: the HadCM3
935 4xCO(2) forcing experiment revisited. *Climate Dynamics*, **33 (6)**, 817–825.
- 936 Loeb, N., H. Wang, A. Cheng, S. Kato, J. Fasullo, K.-M. Xu, and R. Allan, 2016: Observa-
937 tional constraints on atmospheric and oceanic cross-equatorial heat transports: revisiting the
938 precipitation asymmetry problem in climate models. *Climate Dynamics*, **46**, 3239–3257, doi:
939 10.1007/s00382-015-2766-z.
- 940 Lohmann, U., 2017: Why does knowledge of past aerosol forcing matter for future climate
941 change? *Journal of Geophysical Research: Atmospheres*, **22 (9)**, 5021–5023.

942 Ma, S., and T. Zhou, 2016: Robust Strengthening and Westward Shift of the Tropical Pacific
 943 Walker Circulation during 1979–2012: A Comparison of 7 Sets of Reanalysis Data and 26
 944 CMIP5 Models. *Journal of Climate*, **29**, 3097–3118.

945 Mahlstein, I., G. Hegerl, and S. Solomon, 2012: Emerging local warming signals in observational
 946 data. *Geophysical Research Letters*, **39** (21), L21 711.

947 Manara, V., M. Brunetti, M. Maugeri, A. Sanchez-Lorenzo, A., and M. Wild, 2017: Sunshine
 948 duration and global radiation trends in Italy (1959–2013): To what extent do they agree? *Journal*
 949 *of Geophysical Research: Atmospheres*, **122** (8), 4312–4331.

950 Mann, M., B. Steinman, and S. Miller, 2014: On Forced Temperature Changes, Internal Variability
 951 and the AMO. *Geophysical Research Letters*, **41** (9), 3211–3219.

952 Marini, C., and C. Frankignoul, 2014: An attempt to deconstruct the Atlantic Multidecadal Oscil-
 953 lation. *Journal of Climate*, **43**, 607–625.

954 Marvel, K., R. Pincus, G. Schmidt, and R. Miller, 2018: Internal Variability and Disequilibrium
 955 Confound Estimates of Climate Sensitivity From Observations. *Geophysical Research Letters*,
 956 **45** (3), 1595–1601.

957 Marvel, K., G. Schmidt, R. Miller, and L. Nazarenko, 2016: Implications for climate sensitivity
 958 from the response to individual forcings. *Nature Climate Change*, **6**, 386–389.

959 Mauritsen, T., and R. Pincus, 2017: Committed warming inferred from observations. *Nature Cli-*
 960 *mate Change*, **7**, 652–655.

961 McConnell, J., and R. Edwards, 2008: Coal burning leaves toxic heavy metal legacy in the Arctic.
 962 *Proceedings of the National Academy of Science of USA*, **105** (34), 12,140–12,144.

963 McConnell, J., and Coauthors, 2007: 20th-Century Industrial Black Carbon Emissions Altered
 964 Arctic Climate Forcing. *Science*, **317**, 1381–1384.

965 McGregor, H., and Coauthors, 2015: Robust global ocean cooling trend for the pre-industrial
 966 Common Era. *Nature Geoscience*, **8**, 671–677.

967 McGregor, S., A. Timmermann, M. Stuecker, M. England, M. Merrifield, F.-F. Jin, and
 968 Y. Chikamoto, 2014: Recent Walker circulation strengthening and Pacific cooling amplified
 969 by Atlantic warming. *Nature Climate Change*, **4**, 888–892.

970 Medhaug, I., M. Stolpe, E. Fischer, and R. Knutti, 2017: Reconciling controversies about the
 971 'global warming hiatus'. *Nature*, **545**, 41–47.

972 Meinshausen, M., and Coauthors, 2011: The RCP greenhouse gas concentrations and their exten-
 973 sions from 1765 to 2300. *Climatic Change*, **109**, 213–241, doi:10.1007/s10584-011-0156-z.

974 Meinshausen, M., and Coauthors, 2017: Historical greenhouse gas concentrations for cli-
 975 mate modelling (CMIP6) . *Geoscientific Model Development*, **10**, 2057–2116, doi:10.5194/
 976 gmd-10-2057-2017.

977 Menary, M., and R. Wood, 2018: An anatomy of the projected North Atlantic warming hole in
 978 CMIP5 models. *Climate Dynamics*, **50 (7-8)**, 3063–3080.

979 Millar, R., Z. Nicholls, P. Friedlingstein, and A. M.R., 2017: A modified impulse-response repre-
 980 sentation of the global near-surface air temperature and atmospheric concentration response to
 981 carbon dioxide emissions. *Atmospheric Chemistry and Physics*, **17**, 7213–7228.

982 Morice, C., J. Kennedy, N. Rayner, and P. Jones, 2012: Quantifying uncertainties in global and
 983 regional temperature change using an ensemble of observational estimates: The HadCRUT4
 984 data set. *Journal of Geophysical Research*, **117 (D8)**, D08 101.

- 985 Murphy, L., K. Bellomo, M. Cane, and A. Clement, 2017: The role of historical forcings in
986 simulating the observed Atlantic multidecadal oscillation. *Geophysical Research Letters*, **44** (5),
987 2472–2480.
- 988 Myhre, G., and B. Samset, 2015: Standard climate models radiation codes underestimate black
989 carbon radiative forcing. *Atmospheric Chemistry and Physics*, **15**, 2883–2888.
- 990 Myhre, G., and Coauthors, 2013: *Anthropogenic and Natural Radiative Forcing*. Climate Change
991 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment
992 Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cam-
993 bridge, United Kingdom and New York, NY, USA.
- 994 Myhre, G., and Coauthors, 2017: Multi-model simulations of aerosol and ozone radiative forcing
995 due to anthropogenic emission changes during the period 1990-2015. *Atmospheric Chemistry*
996 *and Physics*, **17**, 2709–2720.
- 997 Nazarenko, L., D. Rind, K. Tsigaridis, A. Del Genio, M. Kelley, and N. Tausnev, 2017: Interactive
998 nature of climate change and aerosol forcing. *Journal of Geophysical Research: Atmospheres*,
999 **122** (6), 3457–3480.
- 1000 Newman, M., and Coauthors, 2016: The Pacific Decadal Oscillation, Revisited. *Journal of Cli-*
1001 *mate*, **29**, 4399–4427.
- 1002 O’Reilly, C., M. Huber, T. Woollings, and L. Zanna, 2016: The signature of low-frequency oceanic
1003 forcing in the Atlantic Multidecadal Oscillation. *Geophysical Research Letters*, **43**, 2810–2818.
- 1004 Otterå, O., M. Bentsen, H. Drange, and L. Suo, 2010: External forcing as a metronome for Atlantic
1005 multidecadal variability. *Nature Geoscience*, **3**, 688–694.

- 1006 Otto, A., and Coauthors, 2013: Energy budget constraints on climate response. *Nature Geoscience*
1007 *Correspondance*, **6**, 415–416.
- 1008 Otto, F., D. Frame, A. Otto, and M. Allen, 2015: Embracing uncertainty in climate change policy.
1009 *Nature Climate Change*, **5**, 917–920.
- 1010 Overland, J., 2016: A difficult Arctic science issue: Midlatitude weather linkages. *Polar Science*,
1011 **10 (3)**, 210–216.
- 1012 Pasini, A., U. Triacca, and A. Attanasio, 2016: Evidence for the role of the Atlantic multidecadal
1013 oscillation and the ocean heat uptake in hiatus prediction. *Theoretical and Applied Climatology*,
1014 **129 (3-4)**, 873–880.
- 1015 Pausata, F., L. Chafik, R. Caballeroa, and D. Battisti, 2015: Impacts of high-latitude volcanic
1016 eruptions on ENSO and AMOC. *Proceedings of the National Academy of Science of USA*,
1017 **112 (45)**, 13,784–13,788.
- 1018 Persad, G., and K. Caldeira, 2018: Divergent global-scale temperature effects from identical
1019 aerosols emitted in different regions. *Nature Communications*, **9 (3289)**, 1–9.
- 1020 Piecuch, C., R. Ponte, C. Little, M. Buckley, and I. Fukumori, 2017: Mechanisms underlying
1021 recent decadal changes in subpolar North Atlantic Ocean heat content. *Journal of Geophysical*
1022 *Research: Oceans*, **122**, 7181–7197.
- 1023 Pinto, J., M. Reyers, and U. Ulbrich, 2011: The variable link between PNA and NAO in observa-
1024 tions and in multi-century CGCM simulations. *Climate Dynamics*, **36 (1-2)**, 337–354.
- 1025 Proistosescu, C., and P. Huybers, 2017: Slow climate mode reconciles historical and model-based
1026 estimates of climate sensitivity. *Science Advances*, **3 (7)**, e1602 821.

- 1027 Rahmstorf, S., J. Box, G. Feulner, M. Mann, A. Robinson, S. Rutherford, and E. Schaffernicht,
1028 2015: Exceptional twentieth-century slowdown in Atlantic Ocean overturning circulation. *Nature Climate Change*, **5**, 475–480.
1029
- 1030 Rahmstorf, S., G. Foster, and N. Cahill, 2017: Global temperature evolution: recent trends and
1031 some pitfalls. *Environmental Research Letters*, **12**, 054 001.
- 1032 Regayre, L., and Coauthors, 2014: Uncertainty in the magnitude of aerosol-cloud radiative forcing
1033 over recent decades. *Geophysical Research Letters*, **41**, 9040–9049.
- 1034 Ribes, A., F. Zwiers, J.-M. Azaïs, and P. Naveau, 2017: A new statistical approach to climate
1035 change detection and attribution. *Climate Dynamics*, **48** (1-2), 367–386.
- 1036 Richardson, M., K. Cowtan, E. Hawkins, and M. Stolpe, 2016: Reconciled climate response esti-
1037 mates from climate models and the energy budget of Earth. *Nature Climate Change*.
- 1038 Roberts, C., and Coauthors, 2013: Atmosphere drives recent interannual variability of the Atlantic
1039 meridional overturning circulation at 26.5N. *Geophysical Research Letters*, **40**, 5164–5170.
- 1040 Rohde, R., and Coauthors, 2012: A New Estimate of the Average Earth Surface Land Temperature
1041 Spanning 1753 to 2011. *Geoinformatics and Geostatistics: An Overview*, **1** (1), 1–7.
- 1042 Rotstayn, L., M. Collier, D. Shindell, and O. Boucher, 2015: Why Does Aerosol Forcing Control
1043 Historical Global-Mean Surface Temperature Change in CMIP5 Models? *Journal of Climate*,
1044 **28**, 6608–6625.
- 1045 Rypdal, K., 2012: Global temperature response to radiative forcing: Solar cycle versus volcanic
1046 eruptions. *Journal of Geophysical Research*, **117** (D6), D06 115.
- 1047 Rypdal, K., 2018: The Life and Death of the Hiatus Consistently Explained. *Preprints*, doi:10.
1048 20944/preprints201805.0372.v1.

- 1049 Santer, B., and Coauthors, 2014: Volcanic contribution to decadal changes in tropospheric tem-
1050 perature. *Nature Geoscience*, **7**, 185–189.
- 1051 Sato, M., J. Hansen, M. McCormick, and J. Pollack, 1993: Stratospheric aerosol optical depth,
1052 1850-1990. *Journal of Geophysical Research*, **98**, 22,987–22,994, doi:10.1029/93JD02553.
- 1053 Scaife, A., and Coauthors, 2014: Skillful long-range prediction of European and North American
1054 winters. *Geophysical Research Letters*, **41** (7), 2514–2519.
- 1055 Schmidt, A., and Coauthors, 2018: Volcanic radiative forcing from 1979-2015. *Geophysical Re-*
1056 *search Letters*, **in review**.
- 1057 Schmidt, G., D. Shindell, and K. Tsigaridis, 2014: Reconciling warming trends. *Nature Geo-*
1058 *science*, **7**, 158–160.
- 1059 Schulz, M., and Coauthors, 2012: Atmospheric Transport and Deposition of Mineral Dust to the
1060 Ocean: Implications for Research Needs. *Environmental Science and Technology*, **46**, 10,390–
1061 10,404.
- 1062 Schurer, A., G. Hegerl, and S. Obrochta, 2015: Determining the likelihood of pauses and surges
1063 in global warming. *Geophysical Research Letters*, **42**, 5974–5982.
- 1064 Schurer, A., S. Tett, and G. Hegerl, 2014: Small influence of solar variability on climate over the
1065 past millennium. *Nature Geoscience*, **7**, 104–108.
- 1066 Serreze, M., and R. Barry, 2011: Processes and impacts of Arctic amplification: A research syn-
1067 thesis. *Global and Planetary Change*, **77**, 85–96.
- 1068 Sévellec, F., A. Fedorov, and W. Liu, 2017: Arctic sea-ice decline weakens the Atlantic Meridional
1069 Overturning Circulation. *Nature Climate Change*, **7**, 604–610.

Shen, L., L. Mickley, E. Leibensperger, and M. Li, 2018: Strong dependence of U.S. summertime air quality on the decadal variability of Atlantic sea surface temperatures. *Geophysical Research Letters*, **44**, 12 527–12 535.

Shindell, D., 2014: Inhomogeneous forcing and transient climate sensitivity. *Nature Climate Change*, **4**, 274–277.

Shindell, D., and Coauthors, 2013: Radiative forcing in the ACCMIP historical and future climate simulations. *Atmospheric Chemistry and Physics*, **12**, 2939–2974.

Shiogama, H., T. Nagashima, T. Yokohata, S. Crooks, and T. Nozawa, 2006: Influence of volcanic activity and changes in solar irradiance on surface air temperatures in the early twentieth century. *Geophysical Research Letters*, **33** (9), L09 702.

Singh, H., G. Hakim, R. Tardif, J. Emile-Geay, and D. Noone, 2018: Insights into Atlantic multidecadal variability using the Last Millennium Reanalysis framework. *Climate of the Past*, **14**, 157–174.

Skeie, R., T. Berntsen, M. Aldrin, M. Holden, and G. Myhre, 2014: A lower and more constrained estimate of climate sensitivity using updated observations and detailed radiative forcing time series. *Earth System Dynamics*, **5**, 139–175.

Smith, S., J. van Ardenne, Z. Klimont, R. Andres, A. Volke, and S. Delgado Arias, 2011: Anthropogenic sulfur dioxide emissions: 1850–2005. *Atmospheric Chemistry and Physics*, **11**, 1101–1116.

Solomon, S., J. Daniel, R. Neely III, J. Vernier, E. Dutton, and L. Thomason, 2011: The Persistently Variable "Background" Stratospheric Aerosol Layer and Global Climate Change. *Science*, **333**, 866–870, doi:10.1126/science.1206027.

- 1092 Stephens, G., M. Hakuba, M. Hawcroft, J. Haywood, A. Behrangi, J. Kay, and P. Webster, 2016:
1093 The Curious Nature of the Hemispheric Symmetry of the Earth's Water and Energy Balances.
1094 *Current Climate Change Reports*, **2**, 135–147, doi:10.1007/s40641-016-0043-9.
- 1095 Stevens, B., 2015: Rethinking the Lower Bound on Aerosol Radiative Forcing. *Journal of Climate*,
1096 **28**, 4794–4819, doi:10.1175/JCLI-D-14-00656.1.
- 1097 Stocker, T., and Coauthors, 2013: *Technical summary. In Climate Change 2013: The Physical*
1098 *Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergov-*
1099 *ernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom
1100 and New York, NY, USA.
- 1101 Stolpe, M., I. Medhaug, and R. Knutti, 2017: Contribution of Atlantic and Pacific multidecadal
1102 variability to twentieth century temperature changes. *Journal of Climate*, **30**, in press.
- 1103 Stott, P., S. Tett, G. Jones, M. Allen, J. Mitchell, and G. Jenkins, 2000: External Control of 20th
1104 Century Temperature by Natural and Anthropogenic Forcings. *Science*, **290**, 2133–2137.
- 1105 Stuecker, M., F.-F. Jin, and A. Timmermann, 2015: El Niño-Southern Oscillation frequency cas-
1106 cade. *Proceedings of the National Academy of Science of USA*, **112 (44)**, 13,490–13,495.
- 1107 Suckling, E., G. van Oldenborgh, J. Eden, and E. Hawkins, 2017: An empirical model for proba-
1108 bilistic decadal prediction: global attribution and regional hindcasts. *Climate Dynamics*, **48 (9-**
1109 **10)**, 3115–3138.
- 1110 Suo, S., O. Otterå, M. Bentsen, Y. Gao, and O. Johannessen, 2013: External forcing of the early
1111 20th century Arctic warming. *Tellus A*, **65 (1)**, 20 578.

1112 Sutton, R., G. D. McCarthy, J. Robson, B. Sinha, A. Archibald, and L. Gray, 2017: Atlantic Multi-
 1113 decadal Variability and the UK ACSIS programme. *Bulletin of the American Meteorological*
 1114 *Society*, **98** (9), in press.

1115 Swingedouw, D., J. Mignot, P. Ortega, M. Khodri, M. Menegoz, C. Cassou, and V. Hanquiez,
 1116 2017: Impact of explosive volcanic eruptions on the main climate variability modes. *Global*
 1117 *and Planetary Change*, **150**, 24–45.

1118 Takahashi, C., and M. Watanabe, 2016: Pacific trade winds accelerated by aerosol forcing over the
 1119 past two decades. *Nature Climate Change*, **6**, 768–772.

1120 Tang, Q., X. Zhang, X. Yang, and J. Francis, 2013: Cold winter extremes in northern continents
 1121 linked to Arctic sea ice loss. *Environmental Research Letters*, **8**, 014 036.

1122 Taylor, K., R. Stouffer, and G. Meehl, 2012: An Overview of CMIP5 and the experiment design.
 1123 *Bulletin of American Meteorological Society*, **93**, 485–498.

1124 Thompson, D., J. Cole, G. Shen, A. Tudhope, and G. Meehl, 2015: Early twentieth-century warm-
 1125 ing linked to tropical Pacific wind strength. *Nature Geoscience*, **8**, 117–121.

1126 Ting, M., Y. Kushnir, R. Seager, and C. Li, 2009: Forced and Internal Twentieth-Century SST
 1127 Trends in the North Atlantic. *Journal of Climate*, **22**, 1469–1481.

1128 Titchner, H., and N. Rayner, 2014: The Met Office Hadley Centre sea ice and sea surface tem-
 1129 perature data set, version 2: 1. Sea ice concentrations. *Journal of Geophysical Research: Atmo-*
 1130 *spheres*, **119**, 2864–2889.

1131 Trenberth, K., 1997: The definition of El Niño. *Bulletin of the American Meteorological Society*,
 1132 **78**, 2771–2777.

1133 Tung, K., and J. Zhou, 2013: Using data to attribute episodes of warming and cooling in in-
 1134 strumental records. *Proceedings of the National Academy of the Sciences of the USA*, **110** (6),
 1135 2058–2063.

1136 Uno, I., and Coauthors, 2009: Asian dust transported one full circuit around the globe. *Nature*
 1137 *Geoscience*, **2**, 557–560.

1138 van Oldenborgh, G., L. te Raa, H. Dijkstra, and S. Philip, 2009: Frequency- or amplitude-
 1139 dependent effects of the Atlantic meridional overturning on the tropical Pacific Ocean. *Ocean*
 1140 *Science*, **5**, 293–301.

1141 Vecchi, G., T. Delworth, and B. Booth, 2017: Climate Science: Origins of Atlantic decadal swings.
 1142 *Nature*, **548**, 284–285.

1143 Vernier, J.-P., and Coauthors, 2011: Major influence of tropical volcanic eruptions on the strato-
 1144 spheric aerosol layer during the last decade. *Geophysical Research Letters*, **38** (12), L12 807,
 1145 doi:10.1029/2011GL047563.

1146 Vincze, M., and I. Jánosi, 2011: Is the Atlantic Multidecadal Oscillation (AMO) a statistical
 1147 phantom? *Nonlinear Processes in Geophysics*, **18**, 469–475.

1148 Wang, K., and Coauthors, 2017: Continuously amplified warming in the Alaskan Arctic: Implica-
 1149 tions for estimating global warming hiatus. *Geophysical Research Letters*, **44** (17), 9029–9038.

1150 Way, R., F. Oliva, and A. Viau, 2017: Underestimated warming of northern Canada in the Berkeley
 1151 Earth temperature product. *International Journal of Climatology*, **37** (4), 1746–1757.

1152 Wilcox, L., E. Highwood, and N. Dunstone, 2013: The influence of anthropogenic aerosol
 1153 on multi-decadal variations of historical global climate. *Environmental Research Letters*, **8**,
 1154 024 033.

- 1155 Wild, M., 2009: Global dimming and brightening: A review. *Journal of Geophysical Research*,
1156 **114 (D10)**, D00D16.
- 1157 Wild, M., A. Ohmura, and K. Makowski, 2007: Impact of global dimming and brightening on
1158 global warming. *Geophysical Research Letters*, **34 (4)**, L04 702.
- 1159 Williams, B., and Coauthors, 2017: North Pacific twentieth century decadal-scale variability is
1160 unique for the past 342 years. *Geophysical Research Letters*, **44 (8)**, 3761–3769.
- 1161 Wilson, R., and D. Hathaway, 2006: On the Relation Between Sunspot Area and Sunspot Number.
1162 Tech. Rep. 214324, Marshall Space Flight Center, Marshall Space Flight Center, Alabama.
- 1163 Wolter, K., and M. Timlin, 1998: Measuring the strength of ENSO events - how does 1997/98
1164 rank? *Weather*, **53**, 315–324.
- 1165 Wyatt, M., S. Kravtsov, and A. Tsonis, 2012: Atlantic Multidecadal Oscillation and Northern
1166 Hemisphere's climate variability. *Climate Dynamics*, **38**, 929–949.
- 1167 Xiang, B., B. Wang, J. Li, M. Zhao, and J.-Y. Lee, 2014: Understanding the Anthropogenically
1168 Forced Change of Equatorial Pacific Trade Winds in Coupled Climate Models. *Journal of Cli-*
1169 *mate*, **27**, 8510–8526.
- 1170 Yan, X., R. Zhang, and T. Knutson, 2017: The role of Atlantic overturning circulation in the
1171 recent decline of Atlantic major hurricane frequency. *Nature Communications*, **8 (1695)**, 1–8,
1172 doi:10.1038/s41467-017-01377-8.
- 1173 Yeager, S., and J. Robson, 2017: Recent Progress in Understanding and Predicting Atlantic
1174 Decadal Climate Variability. *Current Climate Change Reports*, **3 (2)**, 112–127.

1175 Zhang, L., and C. Wang, 2013: Multidecadal North Atlantic sea surface temperature and Atlantic
 1176 meridional overturning circulation variability in CMIP5 historical simulations. *Journal of Geo-*
 1177 *physical Research: Oceans*, **118**, 1–20.

1178 Zhang, R., 2017: On the persistence and coherence of subpolar sea surface temperature and salin-
 1179 ity anomalies associated with the Atlantic multidecadal variability. *Geophysical Research Let-*
 1180 *ters*, **44** (15), 7865–7875.

1181 Zhang, R., R. Sutton, G. Danabasoglu, T. Delworth, W. Kim, J. Robson, and S. Yeager, 2016:
 1182 Comment on "The Atlantic Multidecadal Oscillation without a role for ocean circulation". *Sci-*
 1183 *ence*, **352** (6293), 1527a.

1184 Zhang, R., and Coauthors, 2013: Have Aerosols Caused the Observed Atlantic Multidecadal Vari-
 1185 ability? *Journal of the Atmospheric Sciences*, **70**, 1135–1144.

1186 Zhang, W., G. Vecchi, H. Murakami, G. Villarini, T. Delworth, X. Yang, and L. Jia, 2018: Domi-
 1187 nant Role of Atlantic Multidecadal Oscillation in the Recent Decadal Changes in Western North
 1188 Pacific Tropical Cyclone Activity. *Geophysical Research Letters*, **45** (1), 354–362.

LIST OF FIGURES

- 1190 **Fig. 1.** Global radiative forcing components used in our study (a), decomposition of the four AER
 1191 components including indirect aerosol effects (b), and spatial decomposition of the effective
 1192 and non-effective AER (c). Scaled coastal HadOST (blue) and coastal BE anomalies (red)
 1193 in comparison with 60N-60S HadOST (black) in (d) and the same for coastal ERSSTv4,
 1194 coastal GISTEMP and 60N-60S ERSSTv4 in (e). 59
- 1195 **Fig. 2.** Transient Warming Ratios (TWR) to estimate the TCR adjustment factors for WMGHGs,
 1196 AER and VOL for NHem/SHem (lhs) and Land/Ocean (rhs). *Allforcing* warming contri-
 1197 butions in CMIP5, HadCM3, HadOST and the response model (a, b). WMGHG only (c,
 1198 d), AER only (e, f) and VOL only (g, h) contributions in HadCM3 and the response model.
 1199 Modelled VOL (negative) temperature response is shifted by +10 and +25 years merely for
 1200 better readability. The timeline of volcanic eruptions (scaled radiative forcing) is shown in
 1201 black (g, h). All data are low-pass filtered to remove interannual variability. The boxes at
 1202 the bottom show the inferred (diagnosed) warming ratios (TWRD) for WMGHG and AER
 1203 using the product of the ratios of observed (red) and modelled (orange) *allforcing* TWRs (a,
 1204 b), multiplied by the modelled TWRs (orange) for WMGHG (c, d) and AER (e, f). The es-
 1205 timated warming ratios (TWRE) refer to the simulated response model TWR using TWRD.
 1206 Both values are given in light purple. Only the 30 year period of strongest differential warm-
 1207 ing is used for the central TWR estimates. VOL TWR is only a function of the fast response.
 1208 61
- 1209 **Fig. 3.** Summary panel for all the necessary response model parameter, including their justification.
 1210 Global, hemispheric and Land/Ocean TCR scaling factors for WMGHGs, anthropogenic
 1211 aerosols (AER) and volcanic eruptions (VOL) based on the findings shown in Fig. 1 (top
 1212 box). Forcing response time estimates and sensitivities used in this analysis are provided,
 1213 including their source (bottom box). Colour codes for better readability. The pink labels
 1214 in the lower box refer to the original AER-TWRD. In grey the associated coupling factors.
 1215 Surface temperature trends in HadOST (a), CMIP5 (b) and HadCM3 (c) from 1978-2017. 62
- 1216 **Fig. 4.** Fractional variance (square of the model error) for impulse response model uncertainty
 1217 (green), total radiative forcing uncertainty (blue) and internal variability uncertainty (grey)
 1218 in (a). The 1σ (32-68th percentiles) range is shown. We note that internal variability is
 1219 no response model uncertainty in a strict sense as it is added post-hoc (i.e. onto the cal-
 1220 culated temperature). The peaks in the response model uncertainty coincide with volcanic
 1221 eruptions (e.g. Tambora in 1816) eruption. The Internal variability from selected CMIP5
 1222 piControl runs is contrasted with the unforced residuals from the GMST datasets used in this
 1223 study (b). Observed and modelled timeseries are low-pass filtered with a 30 year smoothing
 1224 radius. The standard error is provided in brackets. 63
- 1225 **Fig. 5.** Illustration of the ENSO influence on our results. In the upper graph in each panel, the ob-
 1226 servations are plotted against the response after adding MEI variability to the time series.
 1227 The lower graph shows the raw impulse response model results against the ENSO-corrected
 1228 suite of observational data. Land (including sea ice grid points) is shown on the upper left
 1229 (brown), NHem on the lower left (red), GMST on the upper right (green), Ocean (excluding
 1230 sea ice grid points) in the centre right (purple), and SHem on the lower right (blue). Ob-
 1231 servations from the HadOST composite (pale grey), Cru4CW (yellow), and BE (black) are
 1232 shown. Explained variances (R^2) are given for non-ENSO corrected, model-adjusted (MEI),
 1233 and observation-adjusted (MEI) (Foster and Rahmstorf 2011) low-pass-filtered correlations.
 1234 The WWII correction factors are applied to both instrumental temperature timeseries in each
 1235 panel (except Land). TCR values associated with alternative response model results are pro-
 1236 vided on the right of each panel (1.2-2.0K). 65

1237	Fig. 6.	Evolution of the response model from forcing and response times as applied in Haustein et al. (2017) (H17), with AER as used in CMIP5 (old AER) and the current version using CEDS AER (new AER). Note that the WWII bias correction is only applied in case of new AER in order to illustrate the impact (no change in Land only). The results are shown for Land (a), NHem (b), Global (c), Ocean (d) and SHem (e). The two dashed lines in the lower graph of each panel indicate the variability of the result as a function of the ECS value applied in the response model. The default value of 3.0K corresponds with our central estimate.	67
1245	Fig. 7.	Unforced residual observed variability. Impulse Response Model (IRM) minus HadOST for NHem (a), Global (b), SHem (c), Land (d), and Ocean (f). HadOST Global as in (b) is compared to CruCW4 and BE Global in (e). A 30 year lowess smooth is added in each plot. The revised AMV index is shown in (a). The Multivariate ENSO Index is added in (b). Note that the rhs y-axis labels for AMV (a) and MEI (b) are different.	68
1250	Fig. 8.	Spatial map of correlation coefficients (R) over time between 1850-2016. Positive correlations in red and negative correlations in black. Annual means are used. (a) Time series of the global response model vs HadOST composite. (b) As (a) but with MEI noise added to the global response model time series. (c) Timeseries of the NHem response model vs HadOST. (d) The improved AMV index (van Oldenborgh et al. 2009) vs HadOST. The AMV/NAVI region is highlighted with a red box. (e) As (c), but with 5 year running means applied to both NHem and HadOST. (f) Combination of (c) and (e) where both regressors are detrended and low-pass filtered with a 5 year running mean. (g) As (d), but with both AMV and HadOST being detrended. (h) As (e) but with 10 year running mean. (j) as (f) but with 10 year running mean. (k) As (d), but with both AMV and HadOST being detrended and low-pass filtered with a 20 year running mean. (m) As (e) but with 20 year running mean. (n) As (f) but with 20 year running mean. SHem area is shown in semi-transparent colours to highlight the NHem region of interest.	69
1263	Fig. B1.	Decadal GMST anomalies for the 20th century Reanalysis, all observational data used in this study including the new Hybrid SST dataset (Cowtan et al. 2017), CMIP5 subset and the NorESM1-M global circulation model. Decade from 1850-59 (top) to 2010-17 (bottom) are shown in each row. All anomalies are given relative to the 1901-2000 baseline period. The 1940-49 decade that is affected by the WWII warm bias is highlighted by the red box. . . .	71

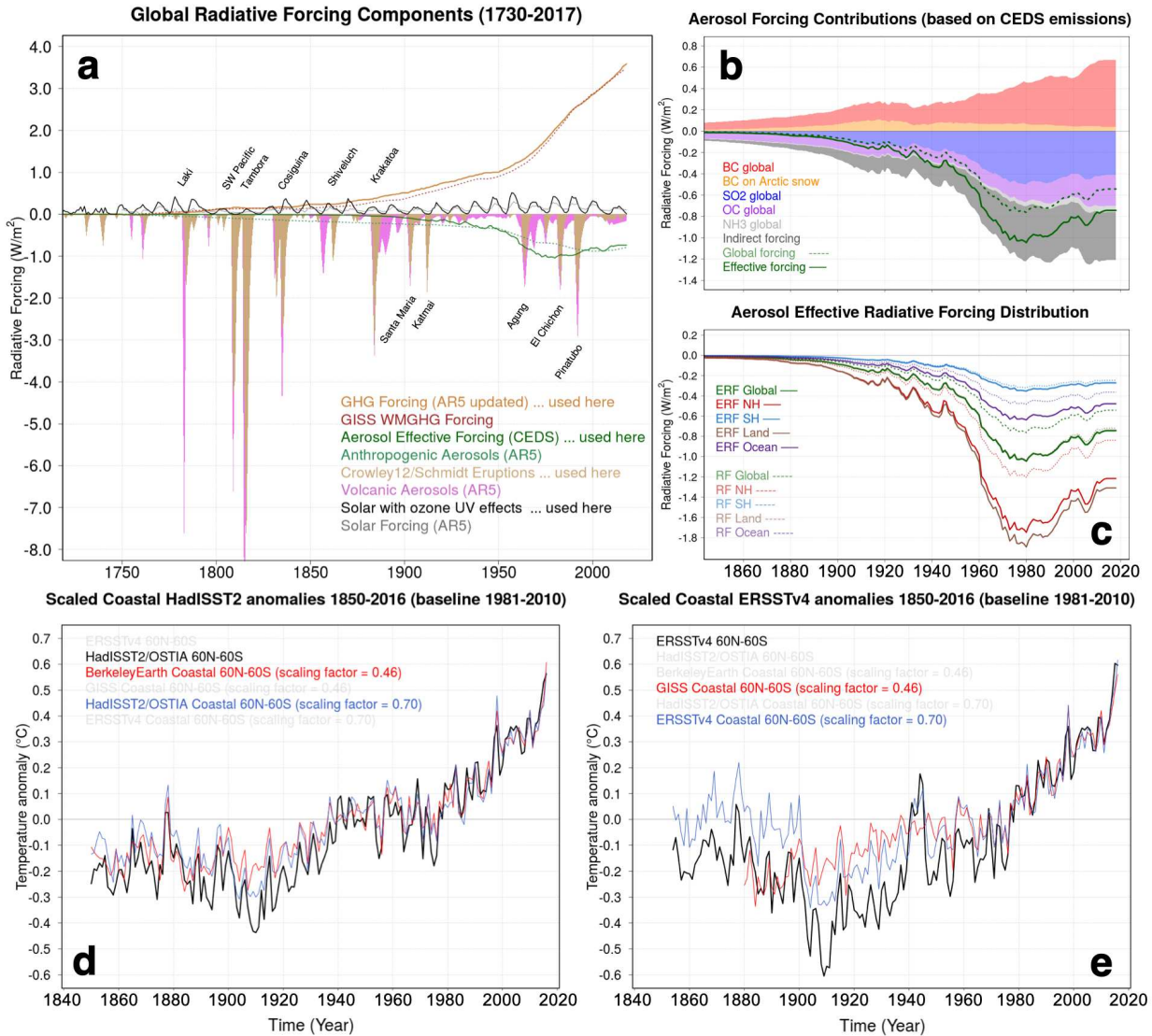
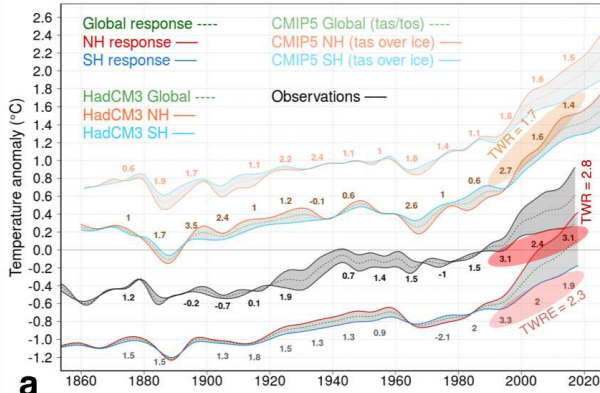


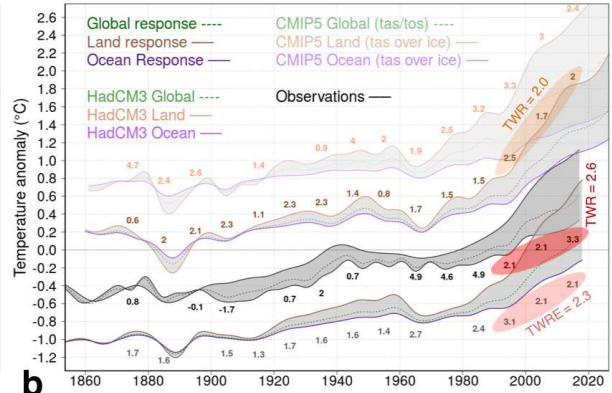
FIG. 1. Global radiative forcing components used in our study (a), decomposition of the four AER components including indirect aerosol effects (b), and spatial decomposition of the effective and non-effective AER (c). Scaled coastal HadOST (blue) and coastal BE anomalies (red) in comparison with 60N-60S HadOST (black) in (d) and the same for coastal ERSSTv4, coastal GISTEMP and 60N-60S ERSSTv4 in (e).

Hemispheric warming contributions: Model vs Observation



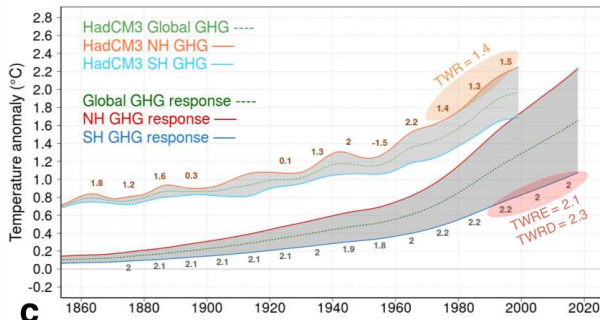
a

Land/Ocean warming contributions: Model vs Observation



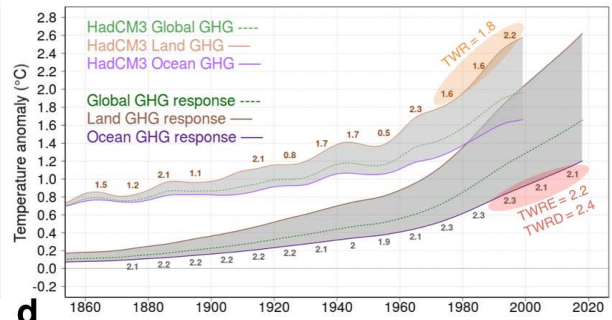
b

GreenhouseGas-Only Forcing: Hemispheric Response Ratio



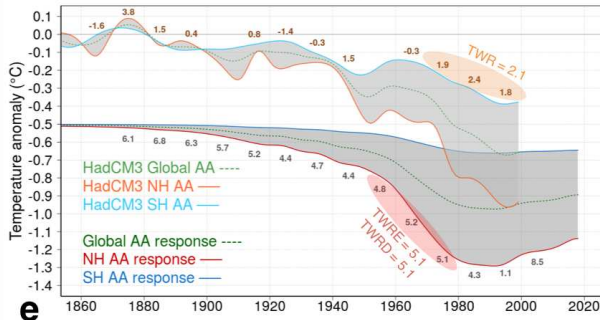
c

GreenhouseGas-Only Forcing: Land/Ocean Response Ratio



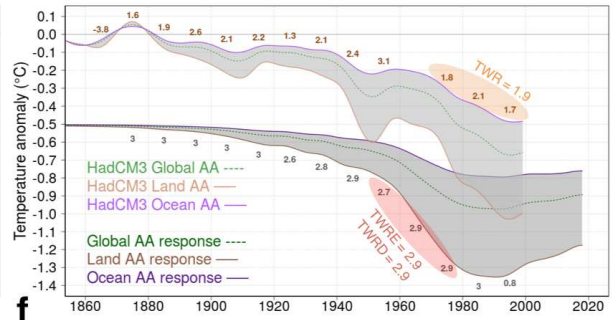
d

Aerosol-Only Forcing: Hemispheric Response Ratio



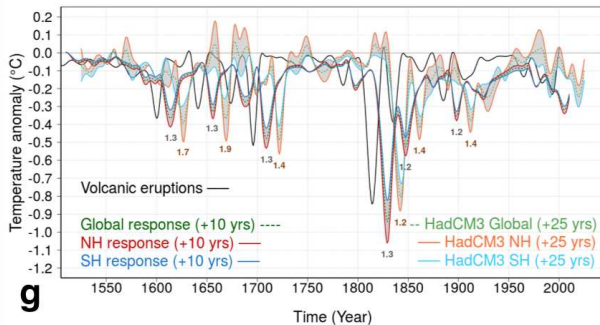
e

Aerosol-Only Forcing: Land/Ocean Response Ratio



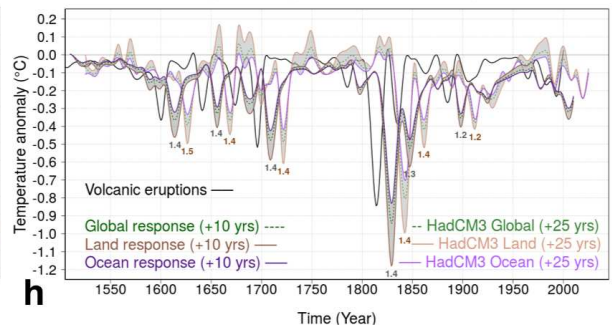
f

Volcanic-Only Forcing: Hemispheric Response Ratio



g

Volcanic-Only Forcing: Land/Ocean Response Ratio



h

TWR = Transient Warming Ratio
TWRE = Estimated Model TWR
TWRD = Diagnosed Model TWR

GHG-Only: NHem/SHem = $1.4 \times (2.8 / 1.7) = 2.3$
AER-Only: NHem/SHem = $2.1 \times (2.8 / 1.7) = 3.5^*$
VOL-Only: NHem/SHem (fast) = $1.3 \times 1.0 = 1.3$

Land/Ocean = $1.8 \times (2.6 / 2.0) = 2.4$
Land/Ocean = $1.9 \times (2.6 / 2.0) = 2.4^{**}$
Land/Ocean (fast) = $1.4 \times 1.0 = 1.4$

1272 FIG. 2. Transient Warming Ratios (TWR) to estimate the TCR adjustment factors for WMGHGs, AER
 1273 and VOL for NHem/SHem (lhs) and Land/Ocean (rhs). *Allforcing* warming contributions in CMIP5, HadCM3,
 1274 HadOST and the response model (a, b). WMGHG only (c, d), AER only (e, f) and VOL only (g, h) contributions
 1275 in HadCM3 and the response model. Modelled VOL (negative) temperature response is shifted by +10 and +25
 1276 years merely for better readability. The timeline of volcanic eruptions (scaled radiative forcing) is shown in
 1277 black (g, h). All data are low-pass filtered to remove interannual variability. The boxes at the bottom show the
 1278 inferred (diagnosed) warming ratios (TWRD) for WMGHG and AER using the product of the ratios of observed
 1279 (red) and modelled (orange) *allforcing* TWRs (a, b), multiplied by the modelled TWRs (orange) for WMGHG
 1280 (c, d) and AER (e, f). The estimated warming ratios (TWRE) refer to the simulated response model TWR using
 1281 TWRD. Both values are given in light purple. Only the 30 year period of strongest differential warming is used
 1282 for the central TWR estimates. VOL TWR is only a function of the fast response.

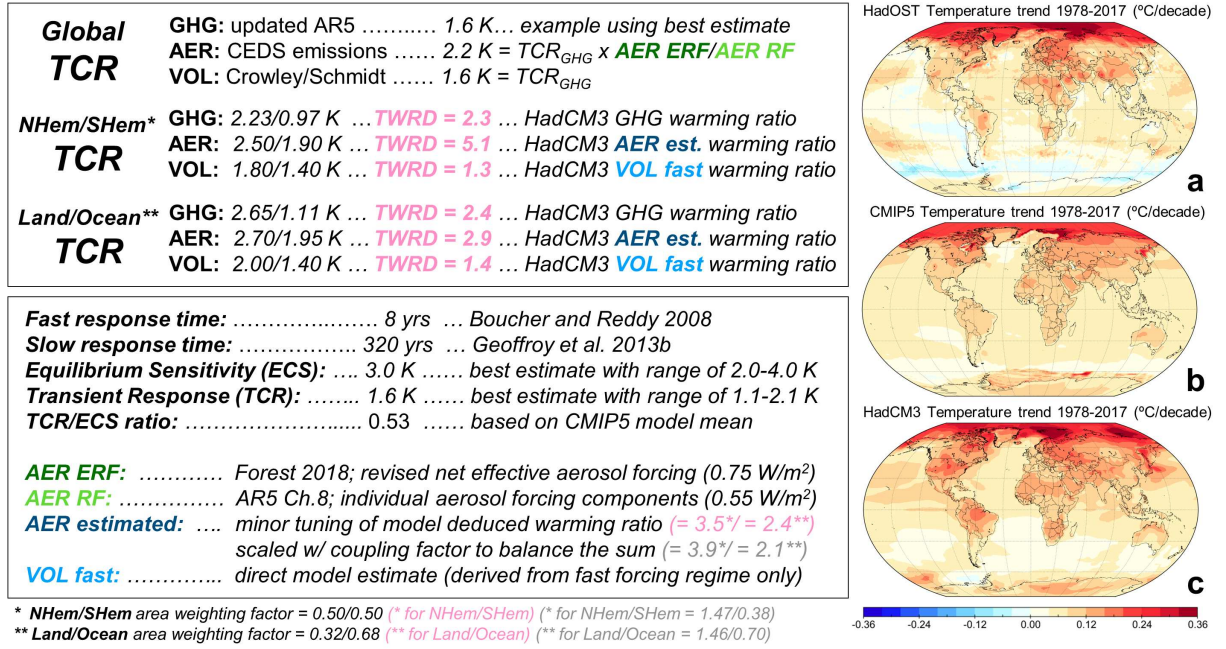


FIG. 3. Summary panel for all the necessary response model parameter, including their justification. Global, hemispheric and Land/Ocean TCR scaling factors for WMGHGs, anthropogenic aerosols (AER) and volcanic eruptions (VOL) based on the findings shown in Fig. 1 (top box). Forcing response time estimates and sensitivities used in this analysis are provided, including their source (bottom box). Colour codes for better readability. The pink labels in the lower box refer to the original AER-TWRD. In grey the associated coupling factors. Surface temperature trends in HadOST (a), CMIP5 (b) and HadCM3 (c) from 1978-2017.

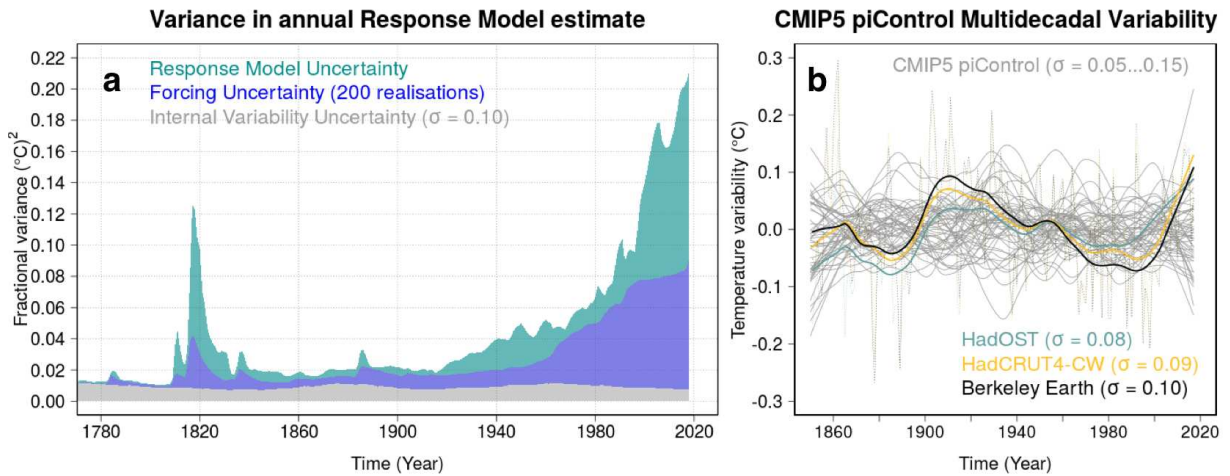


FIG. 4. Fractional variance (square of the model error) for impulse response model uncertainty (green), total radiative forcing uncertainty (blue) and internal variability uncertainty (grey) in (a). The 1σ (32-68th percentiles) range is shown. We note that internal variability is no response model uncertainty in a strict sense as it is added post-hoc (i.e. onto the calculated temperature). The peaks in the response model uncertainty coincide with volcanic eruptions (e.g. Tambora in 1816) eruption. The Internal variability from selected CMIP5 piControl runs is contrasted with the unforced residuals from the GMST datasets used in this study (b). Observed and modelled timeseries are low-pass filtered with a 30 year smoothing radius. The standard error is provided in brackets.

Impulse Response Model vs ENSO/MEI (1850-2017)

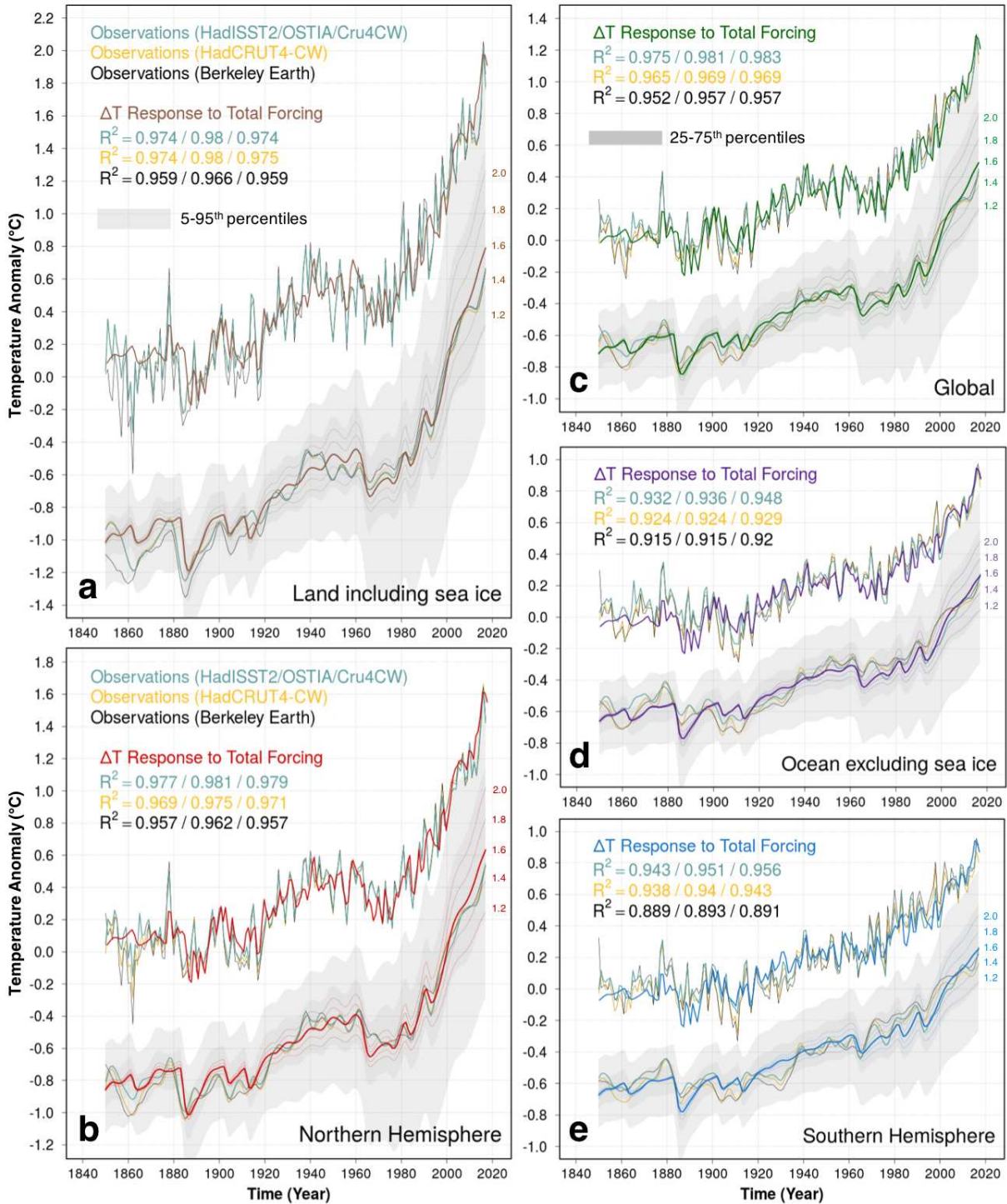
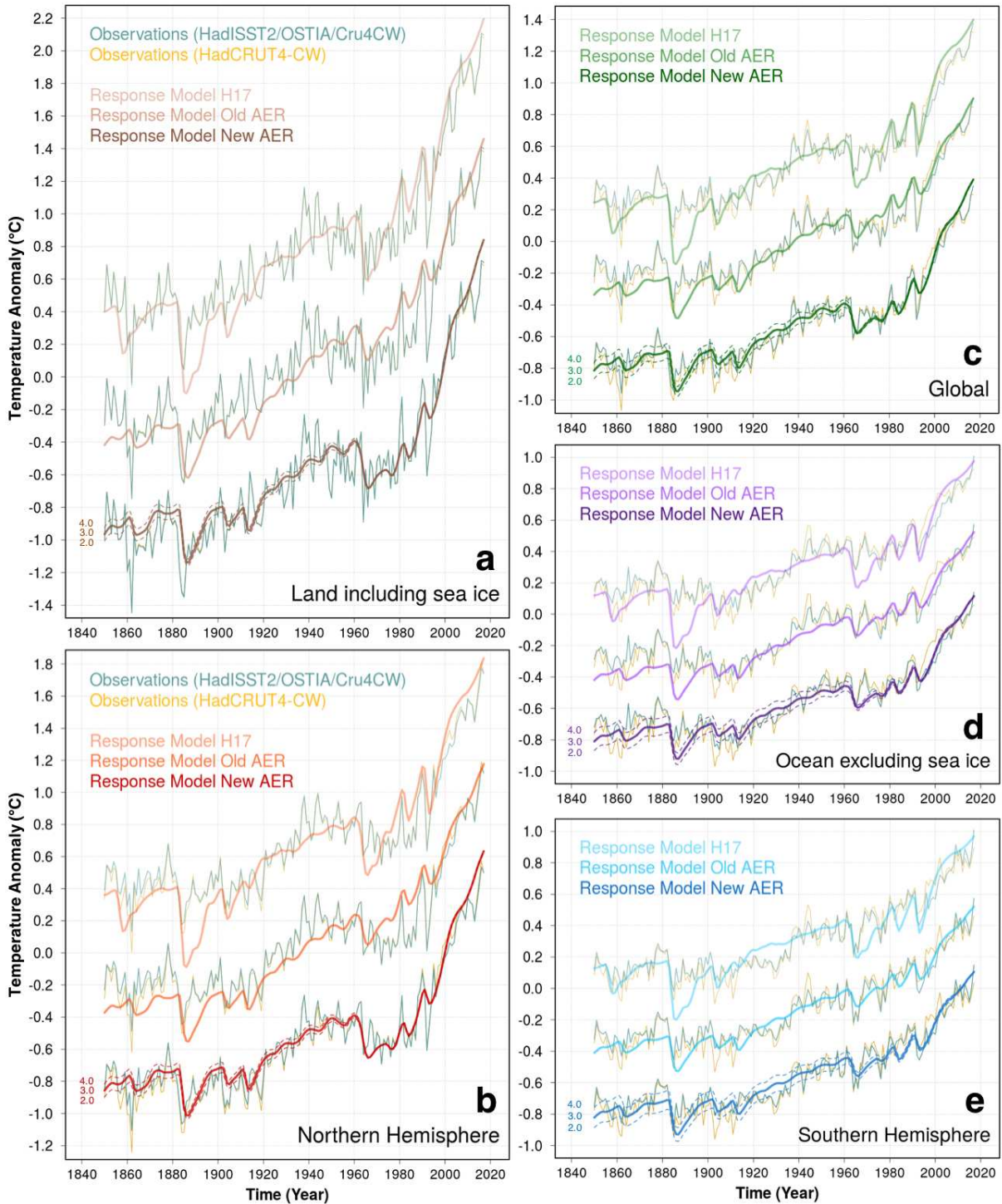


FIG. 5. Illustration of the ENSO influence on our results. In the upper graph in each panel, the observations are plotted against the response after adding MEI variability to the time series. The lower graph shows the raw impulse response model results against the ENSO-corrected suite of observational data. Land (including sea ice grid points) is shown on the upper left (brown), NHem on the lower left (red), GMST on the upper right (green), Ocean (excluding sea ice grid points) in the centre right (purple), and SHem on the lower right (blue). Observations from the HadOST composite (pale grey), Cru4CW (yellow), and BE (black) are shown. Explained variances (R^2) are given for non-ENSO corrected, model-adjusted (MEI), and observation-adjusted (MEI) (Foster and Rahmstorf 2011) low-pass-filtered correlations. The WWII correction factors are applied to both instrumental temperature timeseries in each panel (except Land). TCR values associated with alternative response model results are provided on the right of each panel (1.2-2.0K).

Evolution of the Impulse Response Model



1307 FIG. 6. Evolution of the response model from forcing and response times as applied in Haustein et al. (2017)
1308 (H17), with AER as used in CMIP5 (old AER) and the current version using CEDS AER (new AER). Note that
1309 the WWII bias correction is only applied in case of new AER in order to illustrate the impact (no change in Land
1310 only). The results are shown for Land (a), NHem (b), Global (c), Ocean (d) and SHem (e). The two dashed lines
1311 in the lower graph of each panel indicate the variability of the result as a function of the ECS value applied in
1312 the response model. The default value of 3.0K corresponds with our central estimate.

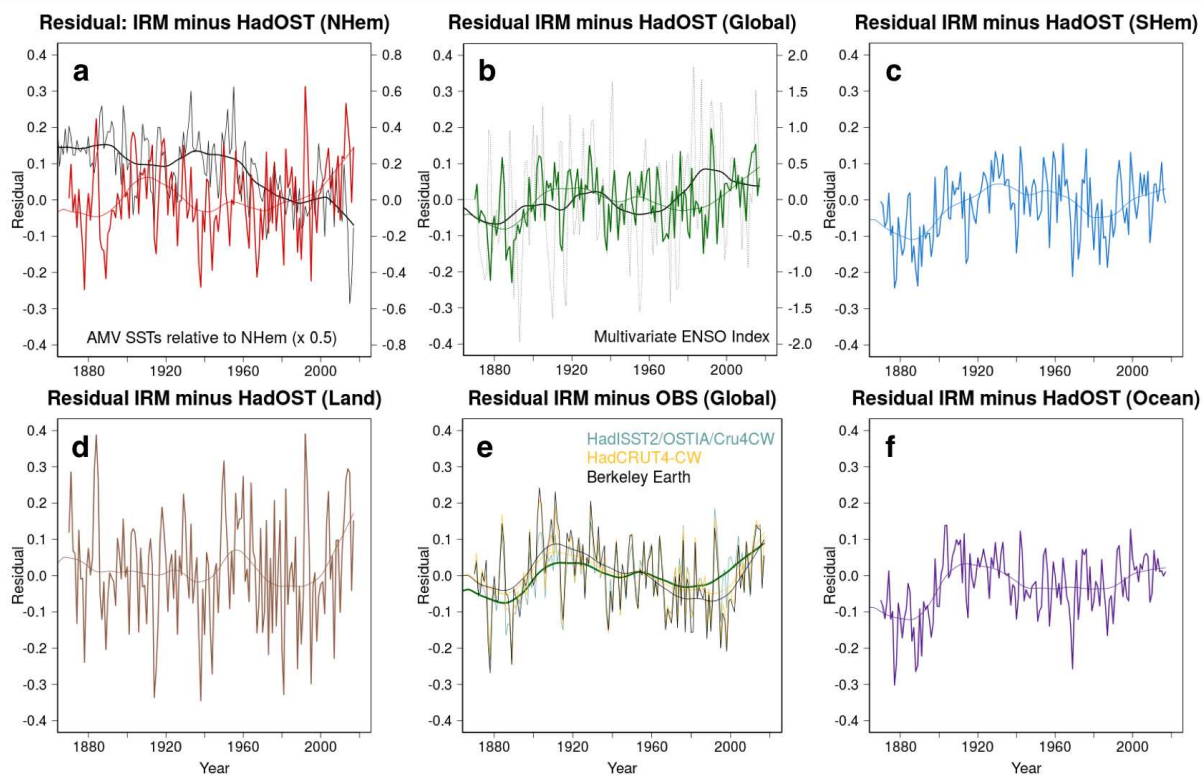


FIG. 7. Unforced residual observed variability. Impulse Response Model (IRM) minus HadOST for NHem (a), Global (b), SHem (c), Land (d), and Ocean (f). HadOST Global as in (b) is compared to CruCW4 and BE Global in (e). A 30 year lowess smooth is added in each plot. The revised AMV index is shown in (a). The Multivariate ENSO Index is added in (b). Note that the rhs y-axis labels for AMV (a) and MEI (b) are different.

Correlation Coefficient between (NHem) Response Model, AMV and HadOST (1850-2016)

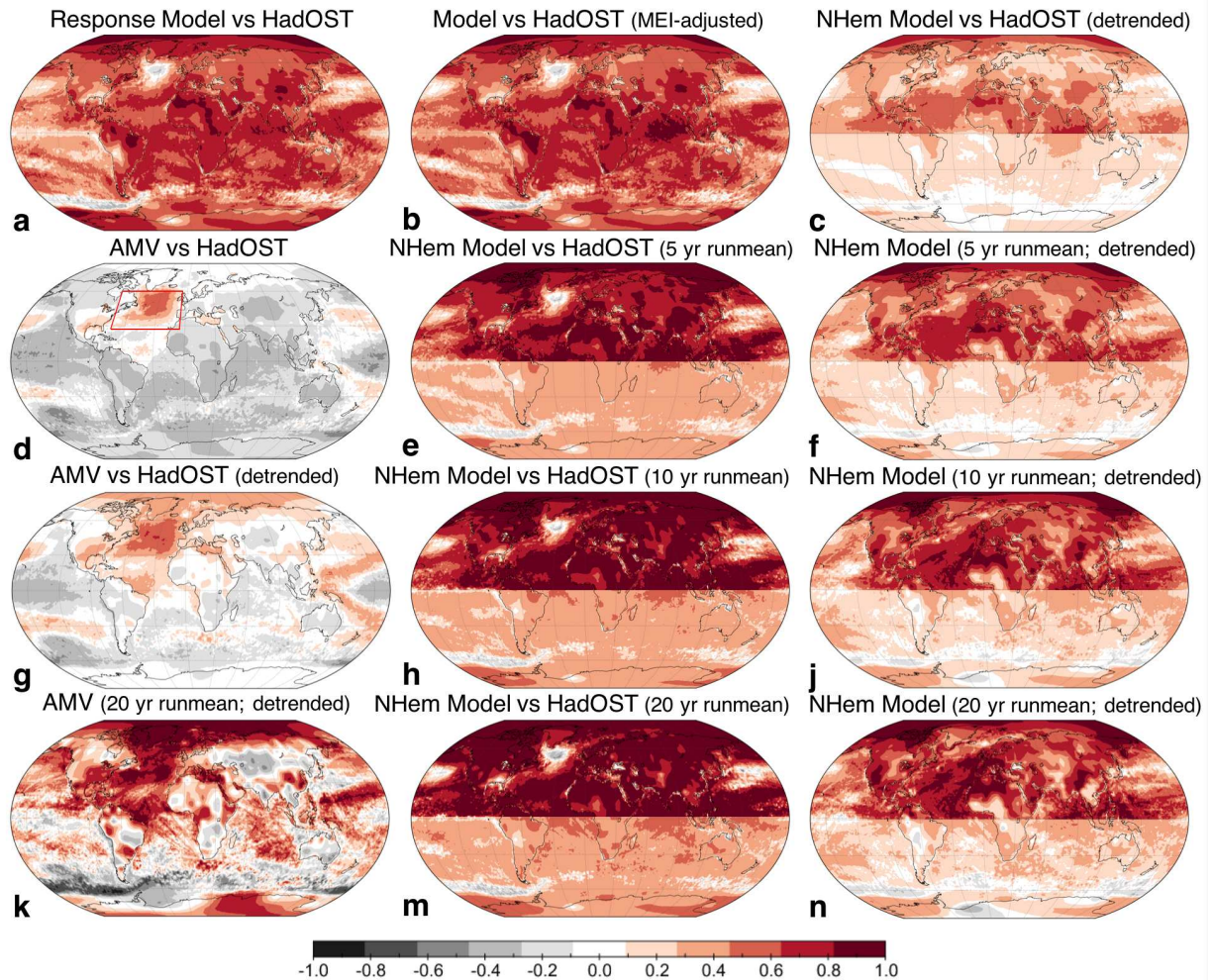
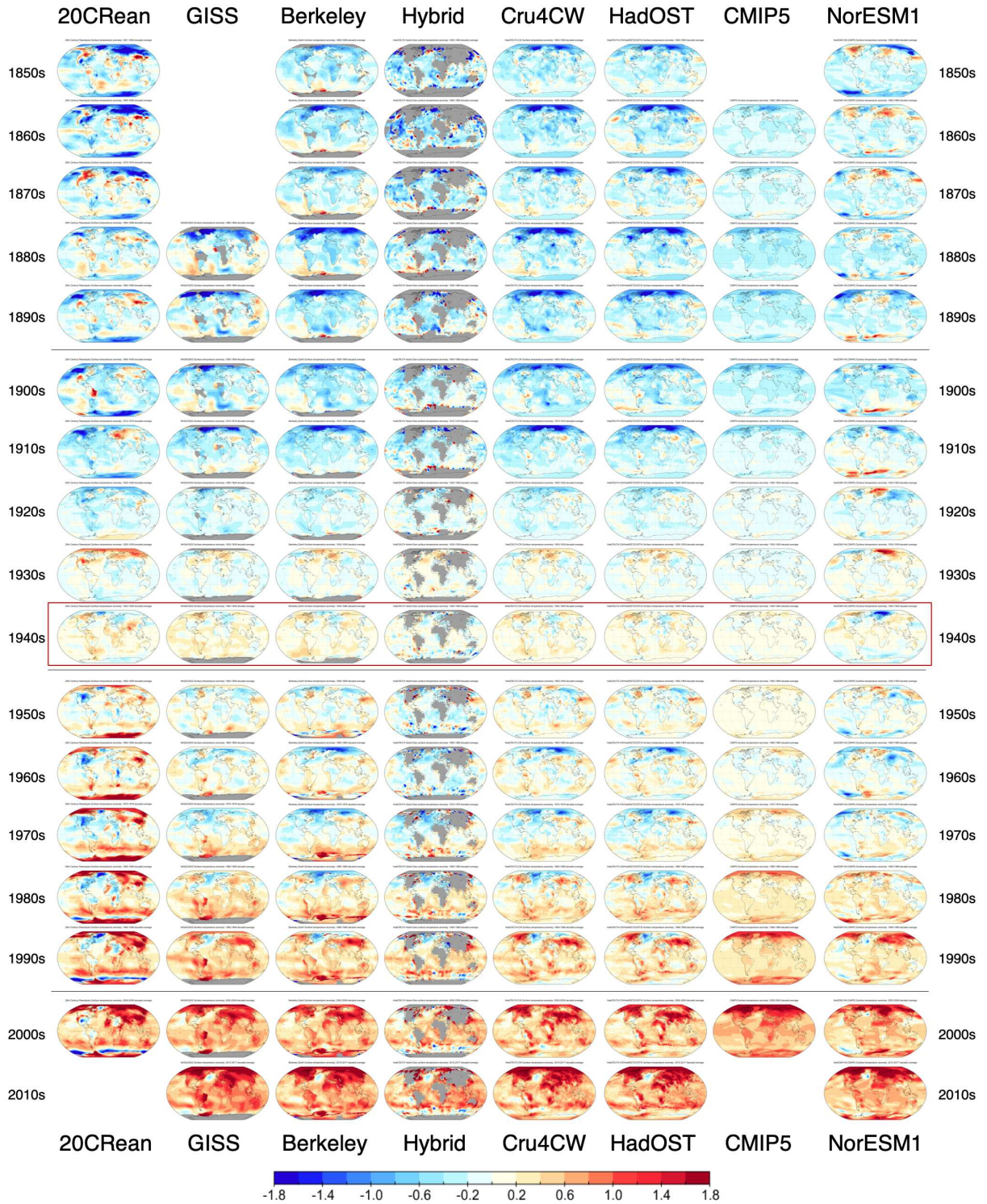


FIG. 8. Spatial map of correlation coefficients (R) over time between 1850-2016. Positive correlations in red and negative correlations in black. Annual means are used. (a) Time series of the global response model vs HadOST composite. (b) As (a) but with MEI noise added to the global response model time series. (c) Timeseries of the NHem response model vs HadOST. (d) The improved AMV index (van Oldenborgh et al. 2009) vs HadOST. The AMV/NAVI region is highlighted with a red box. (e) As (c), but with 5 year running means applied to both NHem and HadOST. (f) Combination of (c) and (e) where both regressors are detrended and low-pass filtered with a 5 year running mean. (g) As (d), but with both AMV and HadOST being detrended. (h) As (e) but with 10 year running mean. (j) as (f) but with 10 year running mean. (k) As (d), but with both AMV and HadOST being detrended and low-pass filtered with a 20 year running mean. (m) As (e) but with 20 year running mean. (n) As (f) but with 20 year running mean. SHem area is shown in semi-transparent colours to highlight the NHem region of interest.

Comparison of decadal air surface temperature averages (1850-2017)



1328 Fig. B1. Decadal GMST anomalies for the 20th century Reanalysis, all observational data used in this
1329 study including the new Hybrid SST dataset (Cowtan et al. 2017), CMIP5 subset and the NorESM1-M global
1330 circulation model. Decade from 1850-59 (top) to 2010-17 (bottom) are shown in each row. All anomalies are
1331 given relative to the 1901-2000 baseline period. The 1940-49 decade that is affected by the WWII warm bias is
1332 highlighted by the red box.